PL-TR-97-2146

THEATER ANALYSIS PROCEDURES (TAP): FINAL REPORT

Thomas Nehrkorn
Ross N. Hoffman
Jeanne Sparrow
Min Yin
Stan Ryckman
Mark Leidner

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139-3794

13 November 1997

Final Report
19 July 1994 - 20 October 1997

PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

This technical report has been reviewed and is approved for publication.

Donald C. Norquist

DONALD C. NORQUIST
Contract Manager

Donald A. Chisholm

DONALD A. CHISHOLM
Acting Branch Chief
Tactical Environmental Support Branch

This report has been reviewed by the ESC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS)

Qualified requesters may obtain additional copies from the Defense Technical Information Center (DTIC). All others should apply to the NTIS.

If your address has changed, if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/VSOS-IM, 29 Randolph Road, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
November 13, 1997

3. REPORT TYPE AND DATES COVERED
FINAL REPORT, 19 July 1994 - 20 October 1997

4. TITLE AND SUBTITLE
Theater Analysis Procedures (TAP): Final Report

5. FUNDING NUMBERS
F19628-94-C-0027
PE 63707F
PR2688 TA GV WU MA

6. AUTHORS: Thomas Nehrkorn, Ross Hoffman, Jeanne Sparrow, Min Yin, Stan Ryckman, Mark Leidner

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Phillips Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010
Contract Monitor: Don Norquist/GPAP

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
PL-TR-97-2146

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report describes the results of a 3-year project for development of a regional analysis system based on the method of optimal interpolation. The TAP project primary objective is to develop robust analysis procedures to support the tactical user. The function of TAP is to combine background (*i.e.* a priori) information with observations of diverse type, quality, and density to produce analyses of meteorological fields, which are either provided directly to the end user or used as input to other algorithms, such as mesoscale forecast models. The prototype system design and its implementation are described, along with the ancillary climatology and error statistics databases. The results of system tests are described in which real data collected over North America were used for the generation of customized analysis products, and for initialization of the MM5 mesoscale forecast model.

14. SUBJECT TERMS
Objective analysis, optimal interpolation, numerical weather prediction, climatology

15. NUMBER OF PAGES
144

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

NSN 7540-01-280-5500

STANDARD FORM 298 (Rev 2-89)
Prescribed by ANSI Std 239-18
298-102

Contents

List of Figures

List of Tables

1 Introduction

This report is the final report of the TAP (Theater-scale Analysis Procedure) project.

The TAP project primary objective is to develop robust analysis procedures to support the tactical user. These analysis procedures provide stable meteorological products for end users.

The function of TAP is to use the optimal interpolation technique to combine background (*i.e. a priori*) information with observations of diverse type, quality, and density to produce analyses of meteorological fields. The TAP analysis configurations are optimized to initialize numerical weather prediction (NWP) models and to provide input for electro-optical tactical decision aids (EOTDAs).

TAP is modular, and capable of utilizing a variety of background and data sources. This capability allows TAP to adapt to different theater meteorological support systems (TMSSs), run on different platforms and to satisfy different user requirements. TAP is configurable to a range of requirements, from first-in stand-alone capability to full Theater Weather Central (TWC) support.

In the nominal case, background fields for TAP are obtained from short-term forecasts of a global NWP model. This requires established communication links, and a set of representative forecast error statistics (standard deviations and correlations). If timely forecasts are not available, older, longer-range forecasts can be used instead, which requires the use of modified error statistics. Finally, if no usable forecast data exist, a climatological background is used.

During the first year of the project (see Nehrkorn *et al.* 1995 [NehHY95]), a preliminary system design was formulated, reviewed by internal and external reviewers, and partially implemented (see Section 2 for a summary of this design). The climatology data required for the background were collected and processed during the first year. A description of the climatology data set is contained in Section 3. To support testing of TAP, real data collection of model output and observational data was begun for selected case studies over the Eastern United States. Finally, an extensive bibliography search was conducted to collect and select appropriate statistics for background and observation errors.

During the second year (see Nehrkorn *et al.* 1996 [NehHS+96]), the system design developed in the first year of the project was implemented in an early prototype system. Most of the placeholders of the preliminary end-to-end system at the end of year 1 were replaced with working code according to the system design. Appendix A, which is a copy of preprint article published in the Proceedings of the Eleventh Numerical Weather Prediction conference, contains a brief system description and illustrations of sample analysis output from the prototype system at that time. A real-data test of the early prototype system was performed by Air Weather Service (AWS) personnel at the Air Force Combat Weather Facility (CWF). This test used real data captured by AER over the first two years of the contract. A description of these real data tests, including the setup and running of the early prototype tests at CWF, and a discussion of the results, is provided in Section 6 and Appendix B.

During the third year, some remaining items of the system design were implemented in the prototype system, and large parts of the prototype code were converted to Fortran 90 to improve its efficiency. The final system development status is fully described in Section 2. System tests were performed to demonstrate the utility of TAP for the initialization of

mesoscale NWP models (Section 7), and prepare the system for an evaluation at the Air Force Weather Agency (AFWA, formerly known as Global Weather Central – GWC; see Section 8 and Appendix C).

The collection of meteorological data for real-data tests of the system, which was started in the first year, was completed in the final year (see Section 5). The required error statistics database established during the first project year was enhanced with additional data and functional fits. Section 4 contains an updated list of the error statistics database, along with plots and tables of some of the error correlations and standard deviations.

2 Prototype System Development

2.1 Functional Description of the TAP Prototype

2.1.1 Information description

For TAP, the OI formalism is configured in several ways, allowing different sets of analysis variables and different analysis grids. The most important OI configurations are the analyses of the meteorological fields to initialize the prediction model. For this purpose the three dimensional fields describing the mass, momentum and humidity of the atmosphere and the two dimensional fields describing the surface conditions are required. TAP follows the approach commonly used in numerical weather prediction and includes configurations of OI for the 3d multivariate analysis of height and wind, the 3d univariate analysis of relative humidity, and the 2d univariate analysis of surface temperature. The 3d multivariate analysis of height and wind uses layer average temperature and surface pressure as well as height and wind information. The univariate analysis techniques are also applicable to arbitrary scalar fields, provided the error statistics can be specified in the same fashion as for relative humidity and surface temperature.

2.1.2 Information flow

In brief OI combines a preexisting background and current observations to yield a quasi-optimal analysis. In addition the OI requires what is essentially a data base of the observational and background error statistics. The background is arbitrary so long as there is a method available to evaluate (ordinarily interpolate) the background to obtain a value analogous to each observed and analyzed datum. The OI formalism treats each observed datum in the same manner. Specificity enters through the observational and background error statistics. The statistical data base may take any form provided a method is available to specify or evaluate the standard deviation of each observed value, the standard deviation of each background value, and the observational error and background error correlations for all observation–observation and observation–analysis value pairs.

2.1.3 Data flow

The TAP involves several procedures, which are grouped into three segments: preprocessing, analysis, and postprocessing. Simply put, the functions of the preprocessing segment are the manipulations needed to produce the input data structures for the OI and to perform preliminary quality control (QC). Similarly the functions of the postprocessing segment are the manipulations needed to optionally spatially smooth the analyzed fields and to extract and reformat the analysis and QC information for the TMSS. Elements of the preprocessing segment which depend on the environment in which TAP is implemented are termed *data ingest* procedures. *Data export* procedures are the analogous elements in the postprocessing segment.

The analysis is invoked several times in different configurations: surface pressure, mass and wind, humidity, and finally surface temperature. TAP is extensible to allow additional

configurations to be added to analyze other variables for different purposes. The TAP overall data flow is shown in Fig. 1.

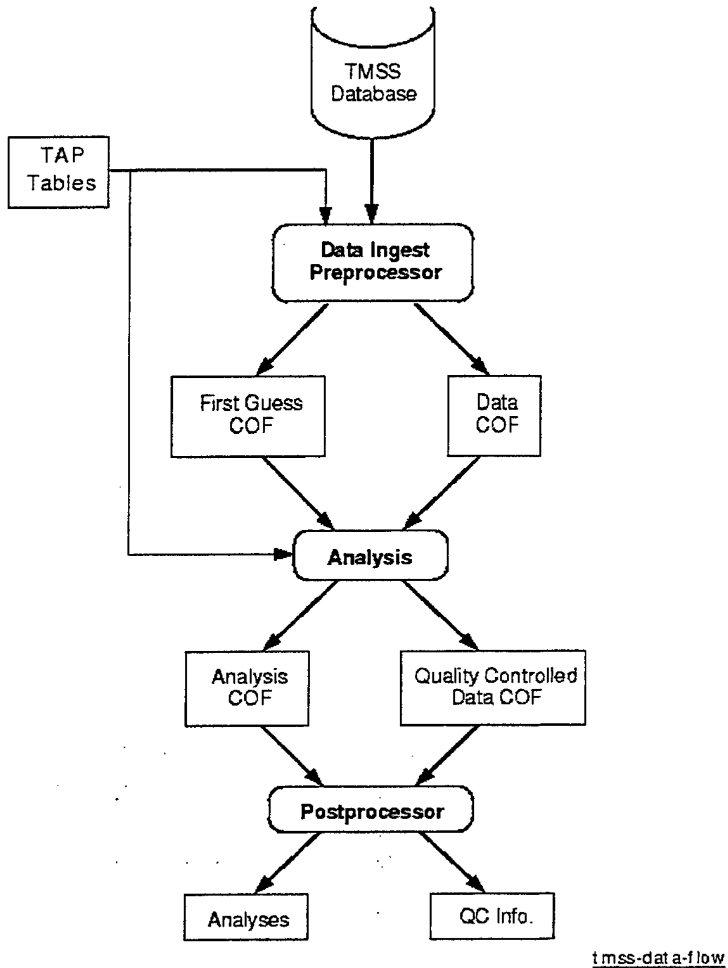


Figure 1: TAP data flow diagram. The Comprehensive Observation Format (COF) is the data format used internally by TAP.

The data preprocessing segment has several main functions. The first is the translation of each specialized data format into the Comprehensive Observation Format (COF). This is the format used internally by TAP. The second generates the output analysis COF file. This sets up the empty COF structure with the proper latitude, longitude, vertical coordinate and variable type for the analysis. The third interpolates the background and, if it is present in the same format, the background error standard deviations for each datum in a COF file. The fourth transforms variables into new variables more suitable for the OI. For example, dew point depression is transformed into relative humidity. The third and fourth preprocessing functions are executed by a single procedure since the background values of several variables are required for the transformation process in some cases. Fifth the data preprocessing provides for each datum an expected standard deviation, if these are not already determined by the COF translation from the data base. Background error standard deviations and error standard deviations due to representativeness and timeliness for both observed and

background values are added at this point. The preprocessor accesses *a priori* statistics specified in tabular form for these purposes. The sixth and seventh data preprocessing functions are QC functions. The first of these, a background QC, is applied to all data. The background QC compares the data to the background, flagging large deviations. The second QC function applies median filter QC and thinning, and optionally super-obbing, to each type of satellite data separately. The median filter provides a "buddy check" and thinning reduces the density of the satellite data to be commensurate with the scales of the analysis. Finally, the preprocessor includes utility functions to merge, sort and select the COF data, as needed, so as to produce a single COF file containing all the data to be used by an analysis.

When preprocessing is complete, the two COF files—the data COF and the analysis COF—contain the proper variables and background values. These files are passed to the analysis segment, which produces two new COF files. The analysis segment performs two main functions, the OI QC and the OI analysis. The new analysis COF file includes the analysis values and optionally the estimated analysis errors. The new data COF file includes QC flags for the data which have been checked. In addition the analysis value and optionally the expected analysis error produced during the OI QC are included for these data.

There are several postprocessing functions needed to produce the desired output data sets from TAP. First, the analysis is transformed into new variables as needed. The QC information from the analysis is reformatted to be passed back to the TMSS data base. Thus a TMSS data base requirement is to allow storage for the various TAP QC flags. The operator optionally performs manual QC by using the TMSS visualization tools to examine the rejected data. If manual QC is performed, a second execution of the TAP is then optionally performed which shows the effect of including or excluding certain observations. Finally, the analysis is regridded. That is, the analyzed fields are reformatted into a more grid-like format to be passed back to the TMSS data base.

The functions described here are all implemented for testing the prototype TAP, but the ultimate TMSS data structures for the various observations and for gridded fields are presently undefined. Consequently, the data ingest and data export requirements cannot be completely specified. The data ingest includes the functions which perform the COF translation and background interpolation. In addition some aspects of the COF generation and variable transformation functions may require some modification depending on the background variables provided by the TMSS and the analysis grid required by the TMSS. The data export includes the reformatting and regridding functions. In addition the smoothing of the analysis might be performed more efficiently for particular grids, but this depends on the analysis grid required by the TMSS. Versions of the data ingest and data export developed during prototyping are necessarily redesigned for operational implementation. For efficiency, we assumed simple and convenient data structures during our prototyping.

Within the TMSS, the data ingest and data export segments interface the COF with the TMSS data structures. The COF files exist only temporarily—the COF make the OI calculations efficient, but they are too "verbose" to be used as the overall TMSS data archive. Generally, the OI data structures are used both for disk and memory storage.

2.1.4 Functional description

Here we provide a technical description of the optimal interpolation (OI) approach to meteorological data analysis and show how it meets the overall TAP requirements.

The OI methodology may be traced back to work by Eliassen (1981, originally published in 1954 [Eli81]) and Gandin (1963 [Gan63]). The first quasi-operational and operational implementations were by Rutherford (1972 [Rut72]) and Schlatter (1975 [Sch75]). The NMC system further evolved from the work of Schlatter as described by Dey and Morone (1985 [DeyM85]), DiMego (1988 [DiM88]), Deaven et al. (1990 [DeaDB+90]), and DiMego *et al.* (1992 [DiMMP+92]). The AF PL global analysis system developed by Norquist (1988 [Nor88]) is based on the NMC system. The NOAA Forecast Systems Laboratory (Benjamin 1989; Benjamin *et al.* 1991; Miller and Benjamin 1992 [Ben89, BenBB+91, MilB92]) has developed a regional analysis system with short update cycles, which makes use of nonstandard data sources such as profiler winds and automated aircraft reports. The most comprehensive example of OI is the ECMWF system. We take this system to be the state of the art. The ECMWF assimilation system has evolved since Lorenc's (1981 [Lor81]) basic description with reformulations and extensions described by Shaw *et al.* (1987 [ShaLH+87]), Lönnberg *et al.* (1986 [LonPH86]), Lönnberg (1988 [Lon88]) and Undén (1989 [Und89]). A complete and up to date description of the assimilation system is provided by ECMWF Research Manual No. 1 (Lönnberg *et al.* 1992 [LonSU92]).

The development of the OI formalism is adequately developed in many of the cited works and is not repeated here. The result of this development is that the analysis increment (analysis minus background) is a weighted sum of the observation increments. The weights are determined from a linear equation with coefficients given in terms of the observational and background error statistics. In practice not all observations are used and the data selection for the analysis at any one point is critical to the method. (Variational methods of analysis use all data simultaneously. These methods are currently under development or newly implemented, e.g., Parrish and Derber (1992 [ParD92]).) Furthermore, the error statistics are not precisely known and are based on empirically developed models. Development of such models is an ongoing and time consuming research topic. Accurate interpolation procedures, especially in the vertical coordinate, are applied to the background field to calculate the observational increments and to evaluate the background field on the analysis grid.

2.1.5 Functional partitioning

TAP implements OI in a particularly flexible manner, taking full advantage of the modularity inherent in the OI formalism. TAP adopts many practices from ECMWF, NMC and NRL as well as from the preliminary RAP work of Burgeson *et al.* (1992 [BurRH+92]). In particular, we utilize separate statistics for the forecast errors and observational errors. This is a low risk choice, following the accepted practice at operational centers. Additionally, this approach enhances modularity since statistics for the various background options and observing systems are independent of each other. The specification of the error statistics is kept separate from the actual OI code, by specifying standard deviations and correlation functions as tables to be interpolated. The preprocessing segment associates an expected error with each observed or background value. Correlations are calculated as needed by

the OI. The vertical coordinate for interpolating the statistics is pressure (p), although the analysis points may be specified arbitrarily in latitude, longitude and pressure.

We note that the MAPS project (Benjamin *et al.* 1991 [BenBB+91]) uses moist potential temperature (θ_v) as the vertical coordinate for calculation of correlations. To accommodate this in TAP involves a large change in the statistical data bases, but a relatively minor change to the OI formalism. The required changes include the following: (1) calculating θ_v during preprocessing for each body using background information as necessary; (2) calculating vertical correlations using θ_v instead of p and (3) updating all the statistics, including the horizontal correlations.

The modular approach we take allows for the exploitation of parallelism in the analysis procedures. The principal computation may be done independently for each analysis volume (a single analysis location or a group of neighboring analysis locations). Thus the same final analysis is obtained if we process the entire domain in one pass or in many sections which are later merged. The merge procedure becomes somewhat more complicated if overlapping volumes are used.

In this section we first consider the requirements for the analysis of relative humidity (RH). After describing the conceptual RH analysis, we describe requirements for the other analysis configurations.

The OI analysis of RH determines a correction to the background of RH as a weighted average of the RH observation deviations from the background. The background or first guess might be climatology or another analysis, but is ordinarily a short term forecast. Thus the background is sometimes termed the first guess, the forecast or the prediction. The weights are determined by solving a set of linear least squares equations which involve the error covariances of the forecast and observations. These weights are also used to calculate the expected analysis error.

The analysis procedures needed for the RH analysis are applicable to any scalar field and are extended in certain ways to provide a proper analysis of mass and wind. For example the stepwise regression algorithm is the same for RH, height, T_s, etc. The same algorithm is used to perform the OI quality control and the analysis at a grid point. Another example is the preliminary data selection. The OI first selects all data within a given distance from the analysis point for further consideration. This selection is independent of the analyzed variable. Even the calculation of covariances is generic. First, the required standard deviations are included in the COF data structure. Then, given the variable and observation types, the proper correlation tables are selected. These tables are interpolated to the proper geographic separation or pressures.

With regard to the mass–wind analysis, we follow the ECMWF and other operational centers, in analyzing geopotential height and horizontal wind components together. This approach is termed a multivariate analysis. The motivation for the multivariate analysis of mass and wind is that outside the tropics, the analysis increments are expected to be nearly geostrophic, and this fact can be used to couple the analyses and use mass and wind data simultaneously. Layer mean temperatures, or thicknesses, can also be used in this analysis. To specify the background error correlations, tables for zz, $z\tilde{u}$, $z\tilde{v}$, $\tilde{u}\tilde{u}$, $\tilde{v}\tilde{v}$, and $\tilde{u}\tilde{v}$, are needed, where z denotes geopotential, and $\tilde{u}$ and $\tilde{v}$ are the wind components in the natural coordinate system defined by the location of the two points in question. By expressing $\Delta z = z_1 - z_2$, four additional correlations involving thicknesses are determined

from the tables already mentioned. Deriving the wind correlations near the pole can be tricky. We use the method of Lorenc (1981 [Lor81]), which first calculates the correlations in the natural coordinate system and then transforms the correlations to the usual wind component system. In this natural coordinate system, the various correlations are modeled in terms of pressure and distance only. This method is ideal for TAP because it is equally applicable to all locations. For a given analysis domain, TAP chooses the correlations applicable to the tropics or extratropics, or to continental or maritime cases appropriately from the TAP statistical data bases. However because the TAP analysis domain is small the correlations are independent of position within the analysis domain.

Following Daley (1991 [Dal91]) two coefficients are usually included in the relationship between correlations in terms of streamfunction, velocity potential and geopotential height and correlations in terms of (z, u, v). These coefficients, denoted μ and ν describe the extent to which the background errors (and hence the analysis increments) are geostrophic and divergent. Often these coefficients are specified differently in different latitude belts to decouple the height and wind analysis near the equator. For added flexibility, we allow μ and ν to be specified as a table in terms of height above topography. This is optionally used to relax the geostrophic and divergenceless constraint in the boundary layer, where friction is important and there is often sufficient data of both mass and wind to allow the analysis of the ageostrophic divergent flow. Nominal values of these (de)coupling coefficients are provided for tropical and extratropical and for continental and maritime cases.

The analysis of surface temperature (T_s) follows the RH analysis except that all vertical correlations are set to unity and only data at the surface are selected. Various other 2d and 3d scalar quantities are capable of being analyzed in exactly the same way as RH and T_s ; only the error statistics need to be changed.

2.2 System Development Status

The algorithm prototyping during the first year of the TAP project had resulted in a preliminary end-to-end analysis system, which was limited to a scalar analysis, and with parts of the full system design either replaced by Splus library functions or omitted. During the second year, a large number of the missing components were coded, unit tested, integrated, and system tested, mostly using the Splus programming language, to allow for rapid prototyping of algorithms. In preparation for the testing at CWF, a graphical and command line user interface was developed, tested, and documented. These are documented in detail in Appendix B. During the third year, some additional components were implemented in Splus, but most of the system development effort concentrated on improving the efficiency by recoding the computationally intensive parts of the calculation in Fortan 90, and on preparing TAP for real-data tests at the AFWA, where it is to be used for initializing the MM5 mesoscale NWP model. The AFWA implementation is summarized briefly here, and documented in detail in Appendix C.

We note that TAP is designed to be hosted within a work station based meteorological display and analysis system. Since this system is not fully defined, TAP formally begins and ends with the interface described by the Comprehensive Observation Format (COF) data structure. Therefore we developed only limited preprocessor and postprocessor capabilities. Graphical and user interface facilities are likewise limited and have been developed solely

for the purpose of testing TAP. Commercial Off-The-Shelf (COTS) software is used where possible. Nevertheless substantial progress has been made in all these areas as well as in the core TAP algorithms and statistics data bases. Section 4 of this report describes the current TAP statistics data bases and calculations involving the statistical quantities.

As detailed in the Software Requirements Specifications, the preprocessing segment has several main functions. The first is the translation of each specialized data format into the format used internally by TAP, namely the COF for the observations, and a gridded data format for the gridded background field. Splus modules have been written to ingest gridded background fields, either from the TAP climatology datasets, or from forecast fields obtained in GRIB format. The background field can be on any type of vertical coordinate system, provided routines and ancillary fields are provided to compute the pressure from the vertical coordinate. The current Splus code supports pressure, σ (including a surface level), and ETA vertical coordinates. For the AFWA implementation, a Fortran 90 module has been written to ingest the background field in the format used by the MM5 preprocessing suite. Modules have been written (in Splus) to ingest and reformat the following data types, in the formats used by the Air Force Interactive Meteorological System (AIMS) system in our real-data capture: surface observations by synoptic, airport, and ship and buoy reporting sites; radiosonde reports; TOVS retrievals of thickness; aircraft reports of winds, temperature and dewpoint. For the AFWA implementation, a Fortran 90 module has been written to read the AFWA format radiosonde data set. The second preprocessing segment generates the output analysis COF file. This sets up the empty COF structure with the proper latitude, longitude, vertical coordinate and variable type for the analysis. This module has been implemented with the capability to generate an analysis COF from a specified (arbitrary) arrangement of analysis longitude and latitude points, or from a grid on a supported map projection (using a GRIB style grid definition). The analysis levels can be any one of the types of vertical coordinate surfaces that are supported for the background field. The third interpolates the background and, if it is present in the same format, the background error standard deviations for each datum in a COF file. For the AFWA implementation, the analysis grid is determined from the gridded background field, and these analysis COF preprocessor functions are implemented in the Fortran 90 background ingest module. The fourth transforms variables into new variables more suitable for the OI. For example, units are changed to the SI units used throughout TAP, and winds are rotated from the model grid to the East/North coordinate system. Fifth the data preprocessing provides for each datum an expected standard deviation, if these are not already determined by the COF translation from the data base. Background error standard deviations and error standard deviations due to representativeness and timeliness for both observed and background values are added at this point. The preprocessor accesses *a priori* statistics specified in tabular form for these purposes. This function has been implemented with the exception of timeliness errors (backgrounds and observations are all assumed to be timely), but including the derivation of background thickness errors from the height error standard deviations and their vertical correlations. The sixth and seventh data preprocessing functions are QC functions. The background check QC has been implemented as part of the preprocessor, and the median filter/data thinning QC has been implemented as an option for horizontally dense data types (currently: surface observations and satellite retrievals). Finally, the preprocessor includes utility functions to merge, sort and select the COF data, as needed, so as to produce a

single COF file containing all the data to be used by an analysis. These have been fully implemented.

There are four postprocessing functions needed to produce the desired output data sets from TAP. First, the analysis is transformed into new variables as needed. This is presently implemented for the rotation of the wind components. Second, the analysis values are optionally smoothed horizontally. Third, the QC information from the analysis is reformatted to be passed back to the TMSS data base. These two functions have not been implemented in the prototype. Fourth, the analysis is regridded. That is, the analyzed fields are reformatted into a more grid-like format. This is presently implemented by using the same GRIB-style grid definition used for the creation of the analysis COF in the preprocessor, with the same restriction that analysis levels are assumed to be isobaric. For the AFWA implementation, a Fortran 90 module has been written to reformat the TAP analysis output into the format used by the MM5 preprocessing/initialization suite.

Graphics capabilities implemented for the purpose of system diagnostics and the testing at CWF include routines for horizontal contour and shade plots and vector displays of gridded fields, including the capability to overlay map backgrounds; facilities for plotting locations and values or vectors of selected headers or bodies of the analysis or data COF. Top level routines have also been written that combine these facilities for a series of standard plots, using reasonable defaults for a variety of graphical parameters (such as the number of contour levels, the number of digits to include on text displays, etc.).

With regard to the TAP analysis algorithms, for the volume method, we have implemented matrix solvers using the LAPACK library. This part of the algorithm, which transforms the preprocessed analysis and data COF by performing the OI analysis, is fully implemented in Fortran 90. The current version of the code supports three- and two-dimensional analyses of one or more variables, using univariate or multivariate correlations. Analyses are performed in analysis volumes defined in terms of horizontal regions of analysis grid points, using observations from data volumes which encompass the analysis volume. Analysis and data volumes are subdivided according to analysis variable, and, optionally, vertical subregions.

An optional OI QC algorithm is implemented in Splus, using the stepwise regression method (SRM) for the solution of the normal equations.

3 Climatology Database

Climatological fields of mean quantities and their standard deviations are needed within TAP to serve as a background field for the analysis in the case that a short-term forecast is not available. There are several different possible datasets that could be used for defining the mean and standard deviation fields. One set of statistics is available from the National Climatic Data Center (NCDC). They are individual monthly means and selected second moments (variances and cross-products) of winds, height, temperature, and moisture at 9 levels between 1000 hPa and 50 hPa. The statistics were derived from the operational analyses of the National Meteorological Center (NMC). These Climate Diagnostics Data Base (CDDB) datasets are available from the present back to 1978 (1992 for specific humidity and temperature). Long-term climatologies have been computed from the CDDB monthly means for parts of the record: a 10-year climatology is available from the National Center for Atmospheric Research (NCAR), but only for the means of winds, temperature, and geopotential height. A 7-year climatology is also available, which includes the means and second moments of winds, temperature, height, and specific humidity. Because TAP requires statistics for relative humidity, none of these datasets fulfills our requirements. An alternative, readily available dataset was therefore used: a set of monthly means and variances going back over 30 years, based on objectively analyzed radiosonde data. This dataset has been maintained by the Geophysical Fluid Dynamics Laboratory (GFDL) and is described in detail in Oort (1983 [Oor83]). As described in more detail in the next section, it contains information on relative and specific humidity, and has high resolution in the boundary layer (4 levels between 1000 hPa and 850 hPa). Because the analyses do not make use of any first guess or background information, they are most trustworthy over the well-sampled continents, whereas they are based on few data points over large oceanic regions.

3.1 The Oort Radiosonde-Based Dataset

The dataset is described in detail in Oort (1983). The input data used in our processing have the following characteristics:

- Available statistics: Individual monthly means and variances for 1958-1994
- Quantities: u, v (zonal and meridional wind), T (temperature), z (height), q (specific humidity), and RH (relative humidity)
- Grid: 2.5° (latitude) × 5° (longitude)
- Levels (hPa): 1000, 950, 900, 850, 700, 500, 400, 300, 200, 100, 50 (q, RH only to 300 hPa)

3.2 Oort Dataset Processing

3.2.1 Computation of Long-Term Means and Variances

A multi-year climatology was computed from the monthly means in the Oort climatology. All the variables listed above were processed, at all available levels. For each of the 12 months,

the long-term mean of the monthly means, the long-term mean of the monthly variances, and the long-term variance of the monthly means were computed for each variable. When used as a background for an analysis, the expected error standard deviation of the mean field is computed as the square root of the sum of the long-term mean of the monthly variance and the long-term variance of the monthly mean.

Most quantities were computed from a 31 year record (January 1959 - December 1989). Because input data were available for shorter periods, the climatology for all points south of 15° S was computed from January 1964 - December 1989 (26 years), and the climatology of RH from January 1974 - December 1989 (16 years).

Before averaging, known deficiencies of the data were removed. In particular, bad values of v at 1000 hPa for May 1978 at 75° longitude were replaced by interpolation between 70° and 80° longitude before averaging; negative values of positive definite quantities (variances of all fields, means of RH and q) were zeroed before averaging; and values at 180° E and 180° W were averaged.

3.2.2 Horizontal Smoothing and Quality Control

Visual examination of the computed climatology revealed that further quality control was needed. In particular, there were some instances of large isolated maxima or minima in the fields (both the mean and interannual variance fields, indicating the presence of one or a few bad individual monthly means in the original data set). Excessively large interannual variances were also found near the poles for some of the variables.

The solution to both problems was to smooth all fields using a nonlinear running filter in the latitude and longitude directions (using the Splus smooth function). Following the smoothing, nonnegative definite quantities (q, RH, and all variances) were reset to be ≥ 0 , and RH and the variance of RH were reset to be ≤ 100 and ≤ 10000, respectively. Before smoothing, the interannual variance was limited to no more than the smoothed long-term mean of the monthly variance.

3.3 Sample Fields and Comparison with CDDB Climatology

Selected fields and levels of the smoothed Oort climatology are shown in Figures 2 through 8 for January. For geopotential height, specific humidity, and temperature, shaded plots of the mean are shown with overlaid contours of the standard deviation. In the case of the winds, vector plots of the mean wind field are shown with standard deviations of the vector wind error (computed from the sum of the variances of the wind components). All four plots are shown for 1000 hPa (Figures 2 and 3), 500 hPa (Figures 4 and 5), 300 hPa (Figures 6 and 7). Only the height and winds are shown for 50 hPa (Figure 8).

As is to be expected, the mean fields have a generally smooth appearance, but they clearly contain the salient features of the general circulation: the major baroclinic zone in the tropospheric midlatitudes is evident in the height and temperature maps, and coincides with the belt of strong westerlies. The major regions of cyclogenesis along the east coasts of North America and Asia are associated with a trough in the mean fields. Maxima in the standard deviations are located downwind from the troughs, along the major storm tracks.

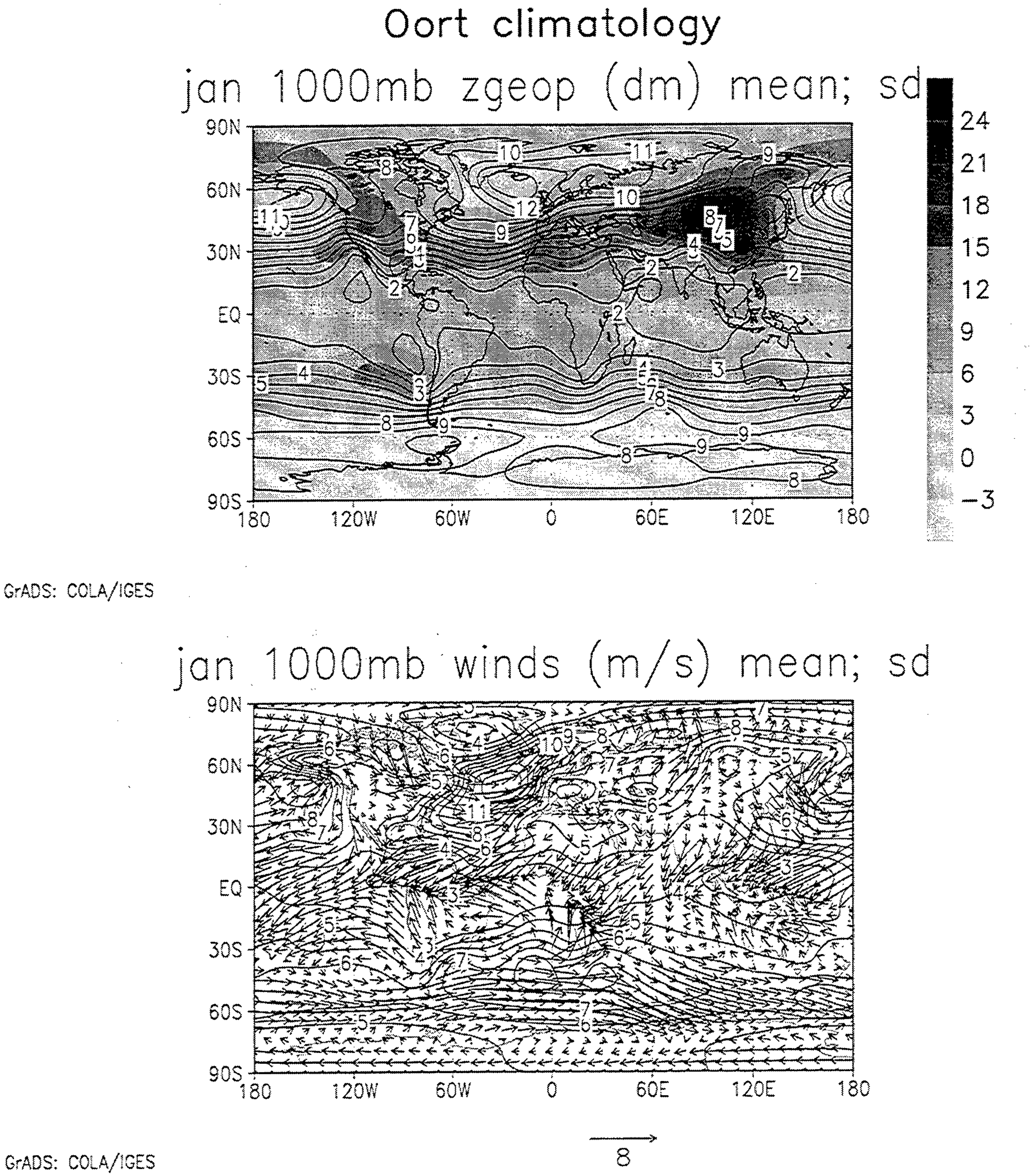


Figure 2: Oort climatology maps for January: 1000 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

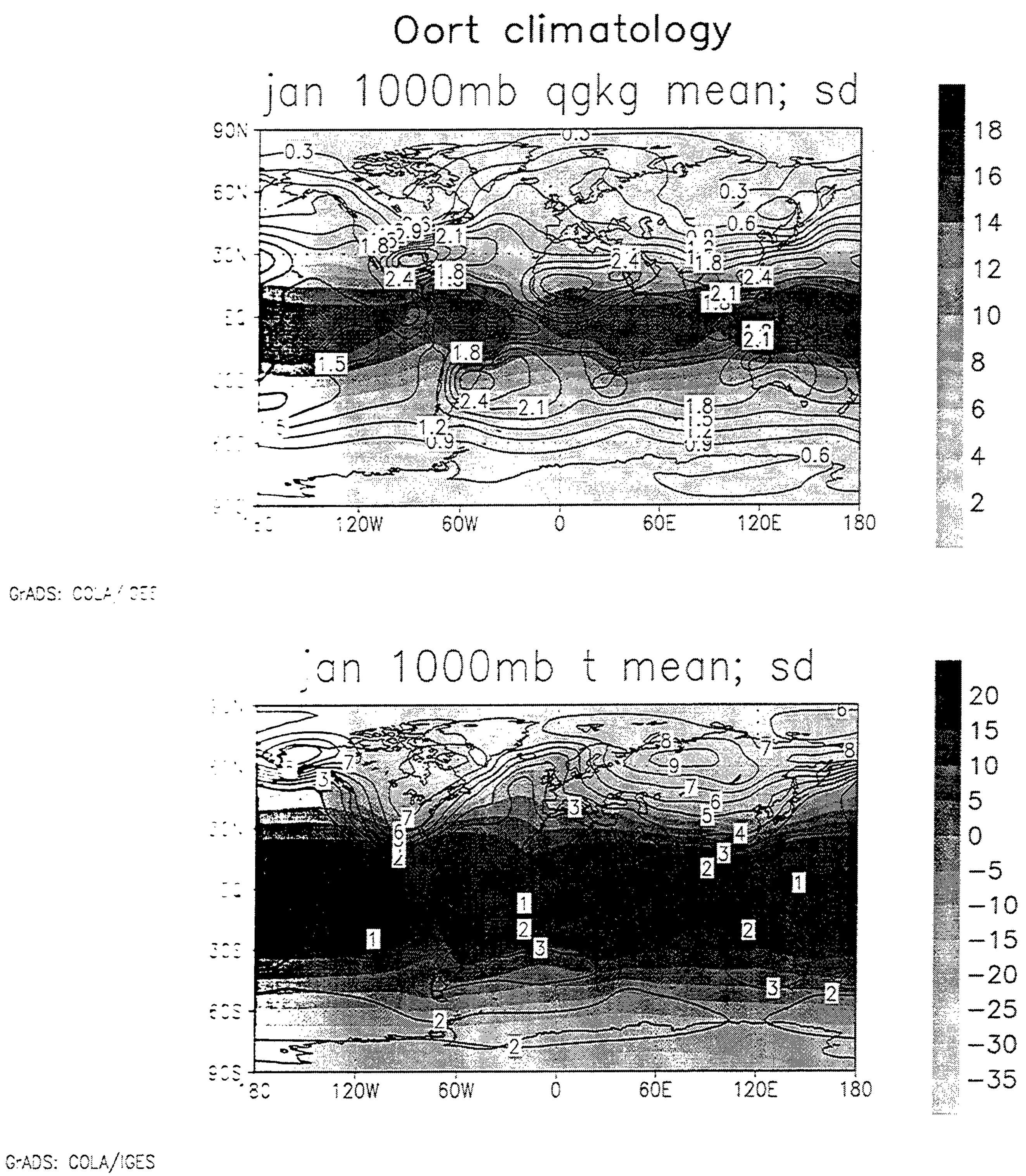


Figure 3: Oort climatology maps for January: 1000 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

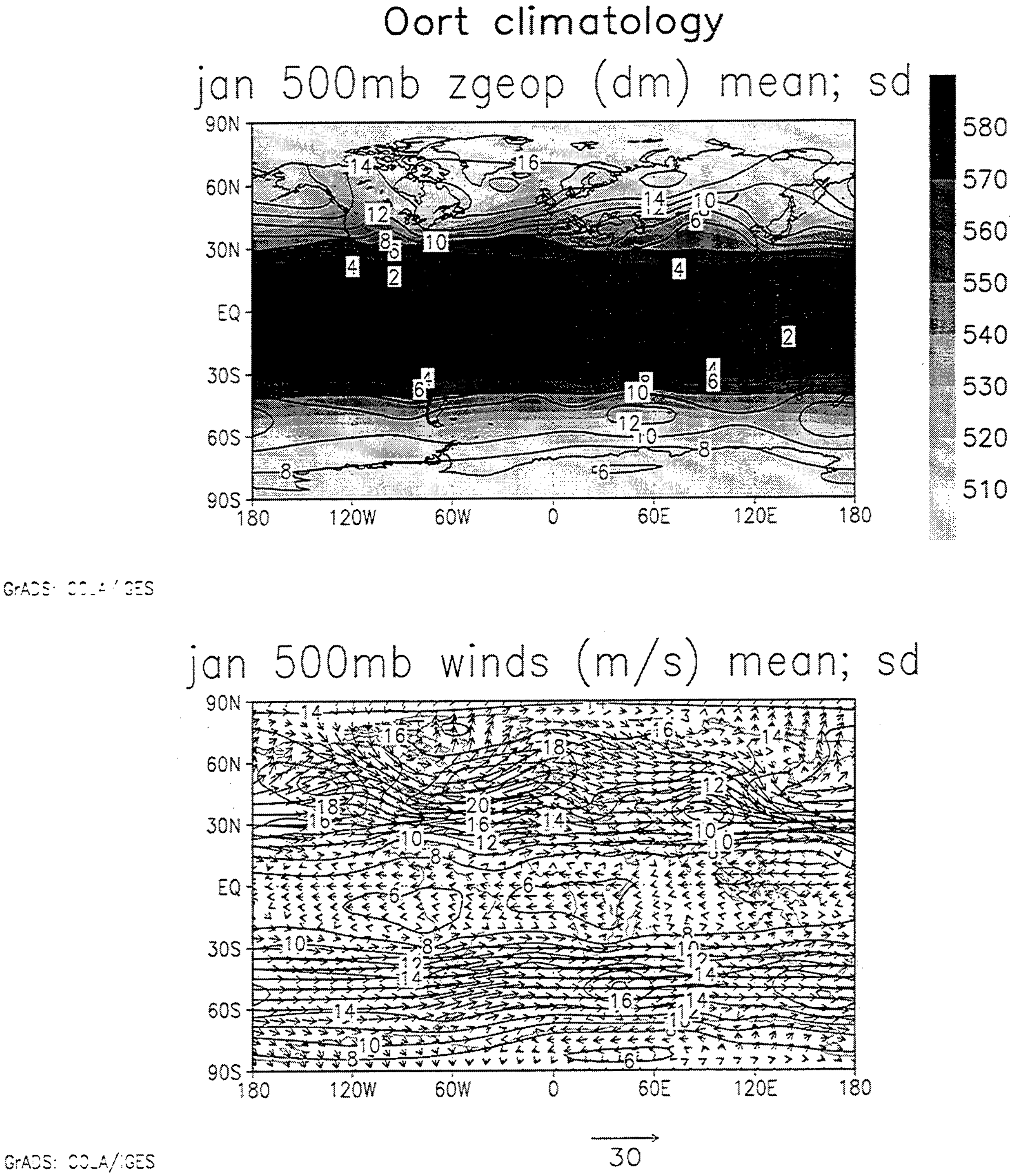


Figure 4: Oort climatology maps for January: 500 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

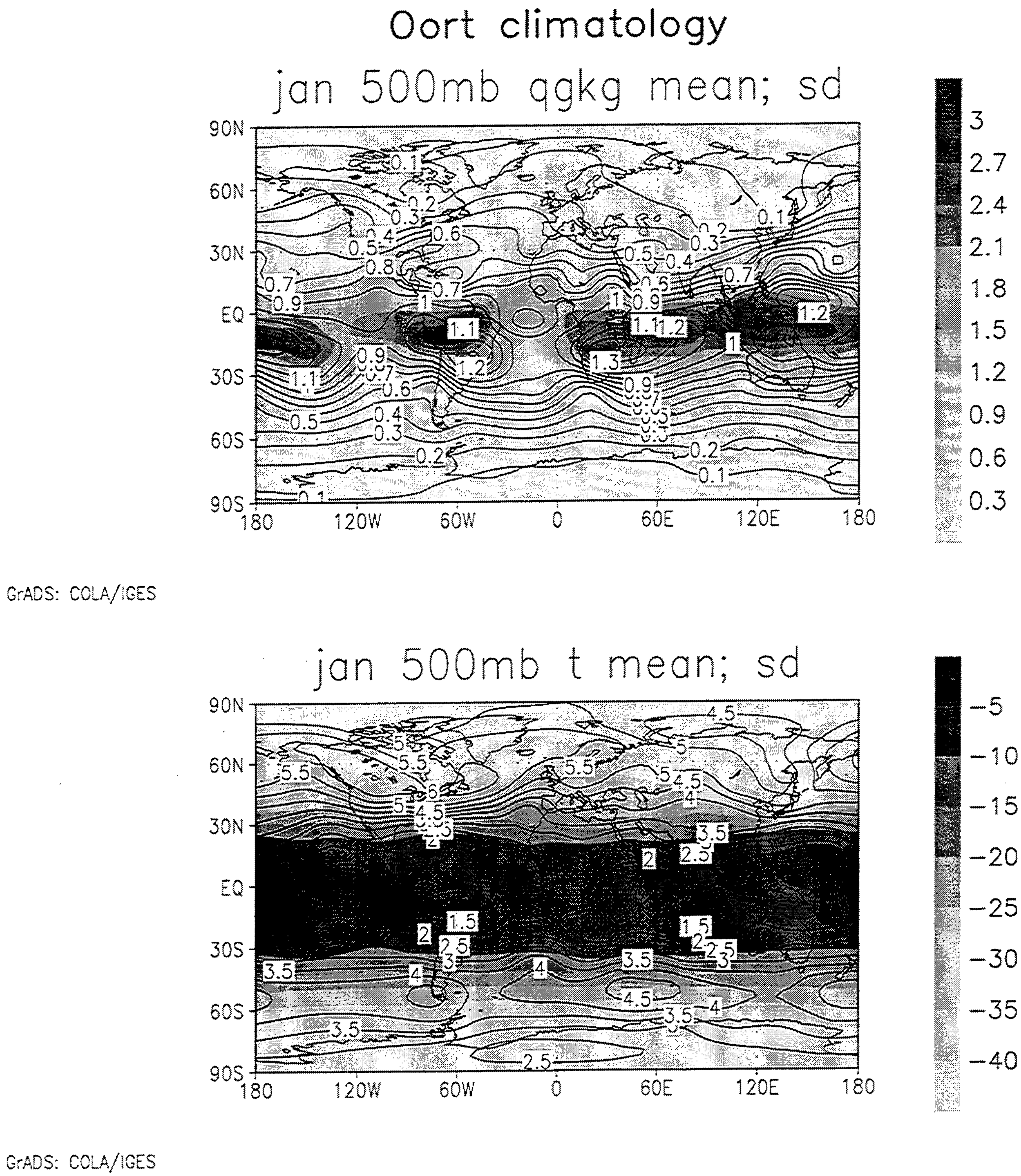


Figure 5: Oort climatology maps for January: 500 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

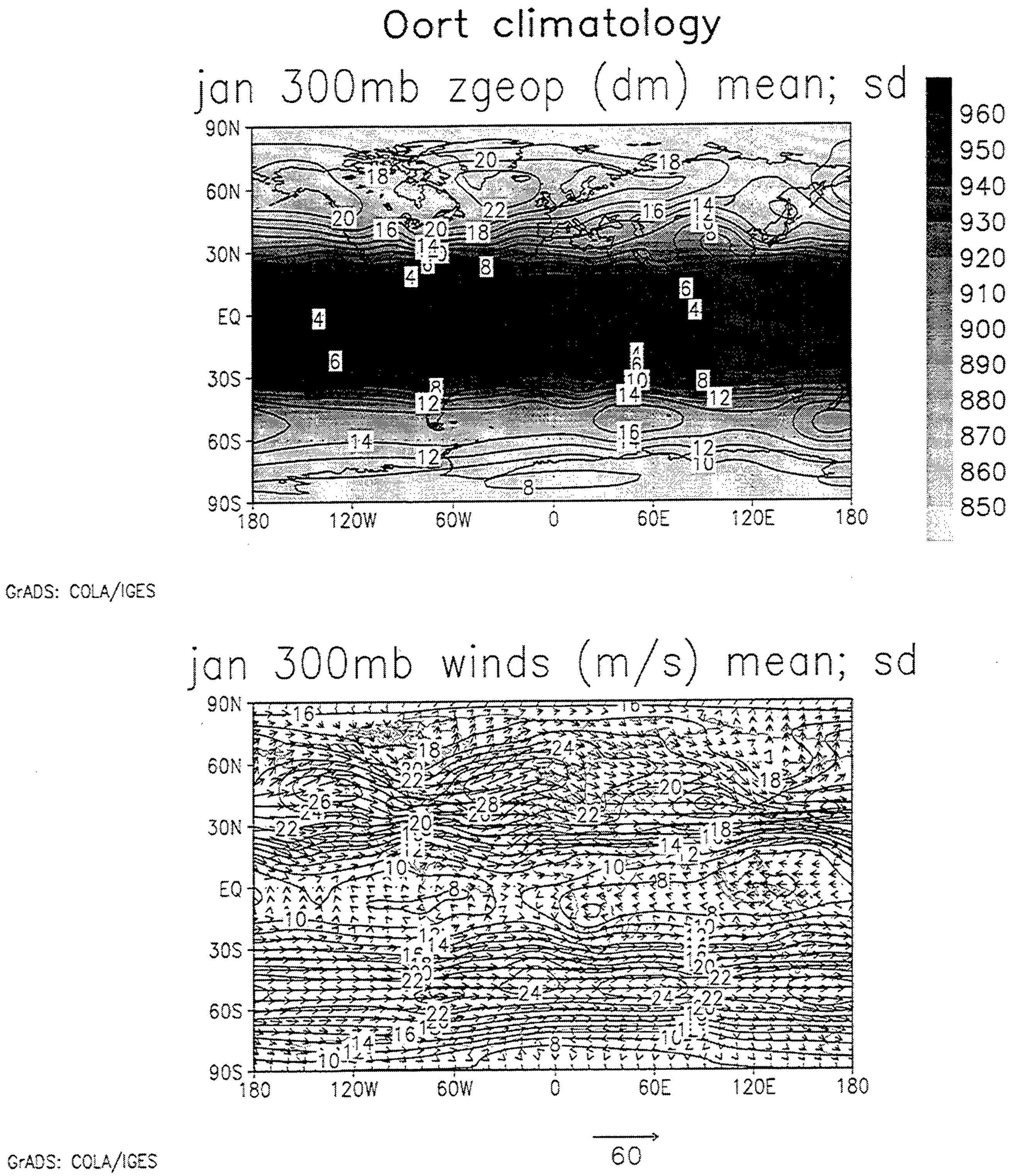


Figure 6: Oort climatology maps for January: 300 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

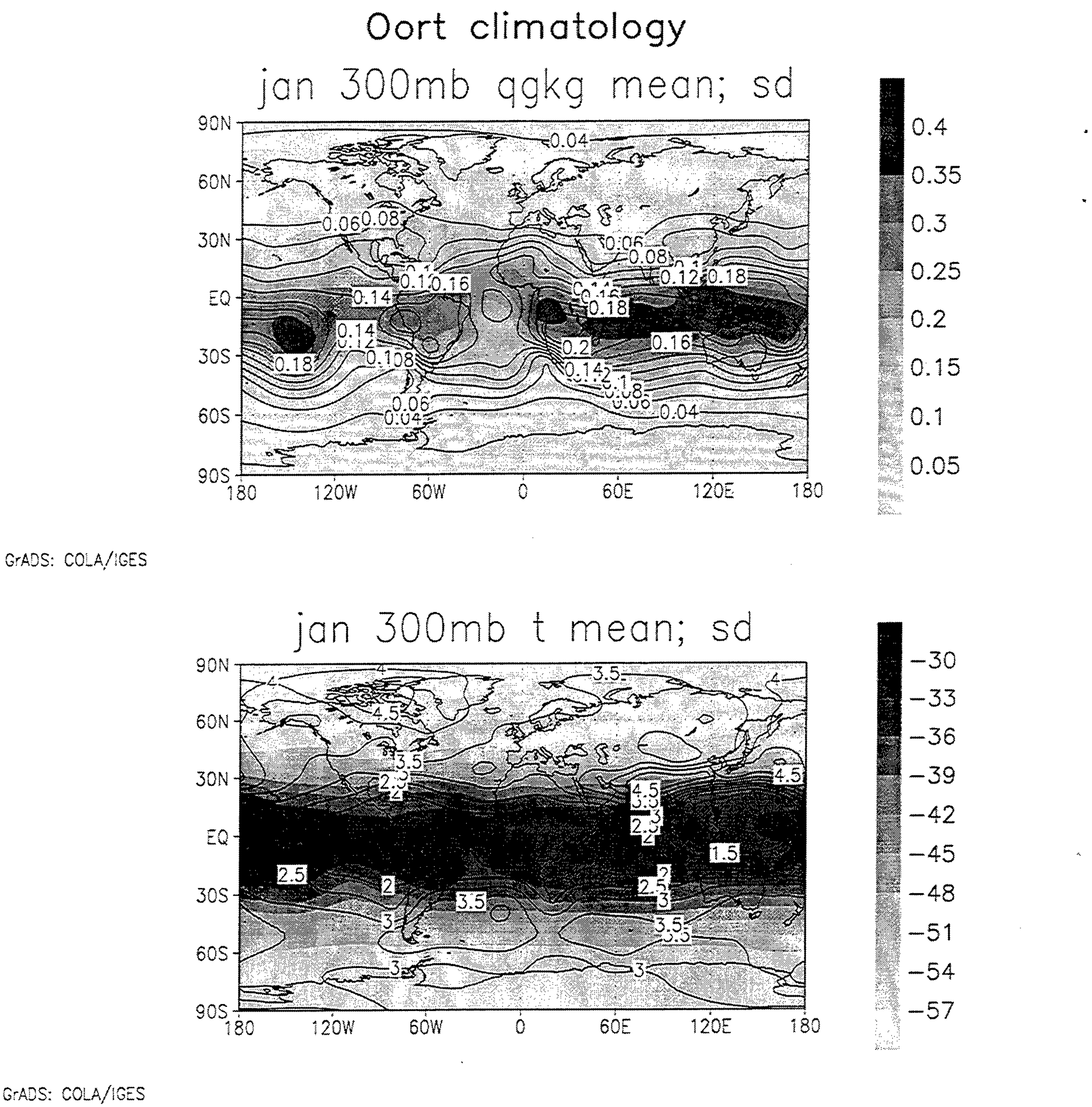


Figure 7: Oort climatology maps for January: 300 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

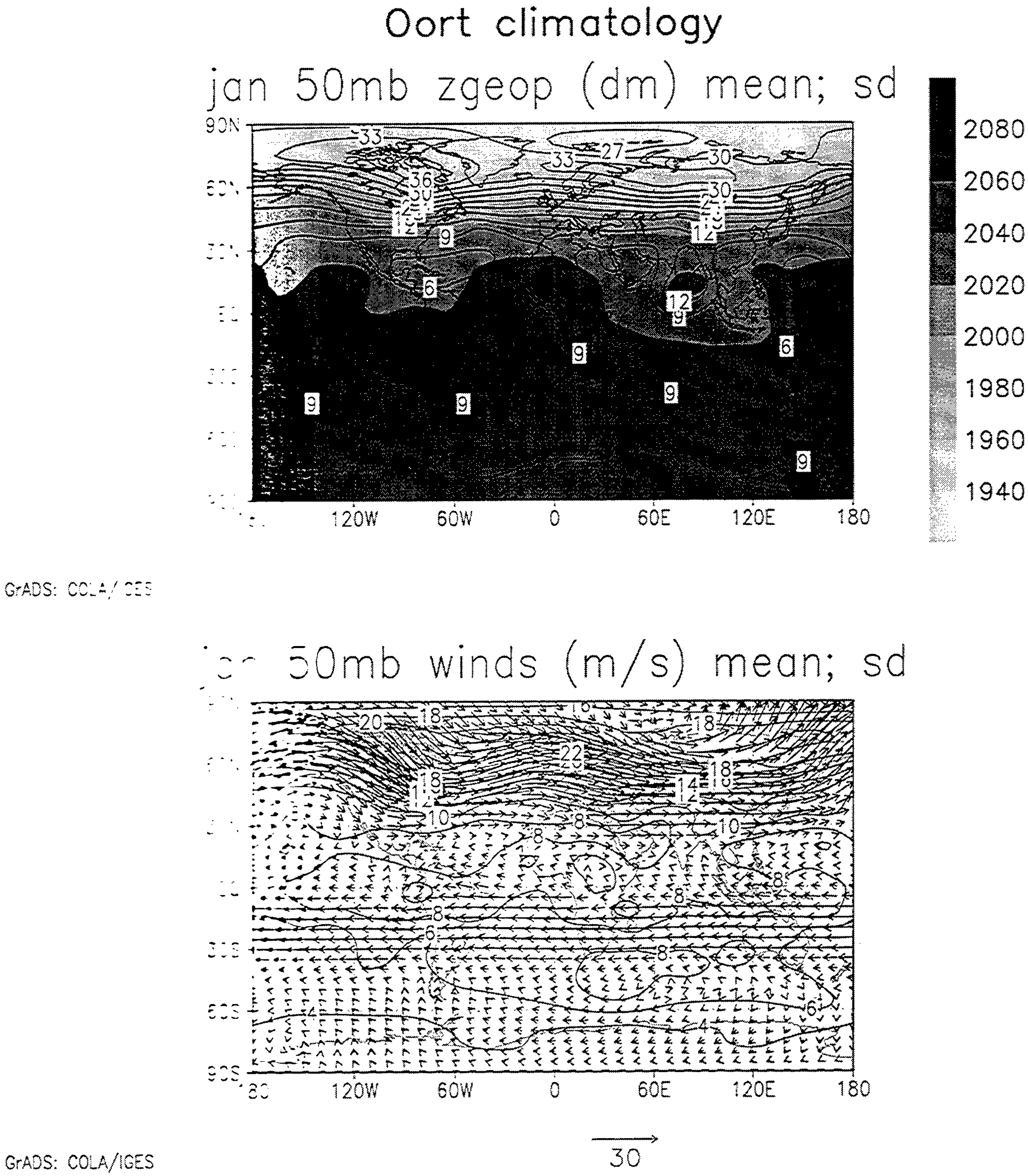


Figure 8: Oort climatology maps for January: 50 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

The corresponding set of figures for July are shown in Figures 9 through 9. They show the typical shift of the strongest baroclinic activity to the Southern Hemisphere. The maps of specific humidity clearly show the effects of the Indian monsoon, in the form of a pronounced maximum of the mean and standard deviation over the area.

The gross features of the mean atmospheric state as defined by the Oort climatology can also be seen in the meridional cross sections of zonal mean temperature and zonal wind shown in Figure 16. The meridional gradient of temperature, the location and strength of the jet stream, and their variation with the seasons all follow the expected pattern.

For purposes of comparison and validation of the climatology dataset computed here, we present horizontal maps for July for the 7-year CDDB climatology dataset obtained from NCAR in Figures 17 through 22. Because moisture and temperature data are not available above 500 hPa in this dataset, these fields are not shown at 300 hPa. Comparison with Figures 9 through 15 shows similarities in all the major circulation features. The effects of the smoothing of the Oort climatology are evident by the somewhat noisier appearance of the mean and particularly the standard deviation fields in the CDDB data. The meridional cross sections shown in Figure 23 exhibit even fewer differences from the corresponding Oort climatology plots.

In summary, then, both a subjective evaluation and comparison with an independent climatology dataset confirm that the Oort climatology dataset derived here provides reasonable values of the required quantities.

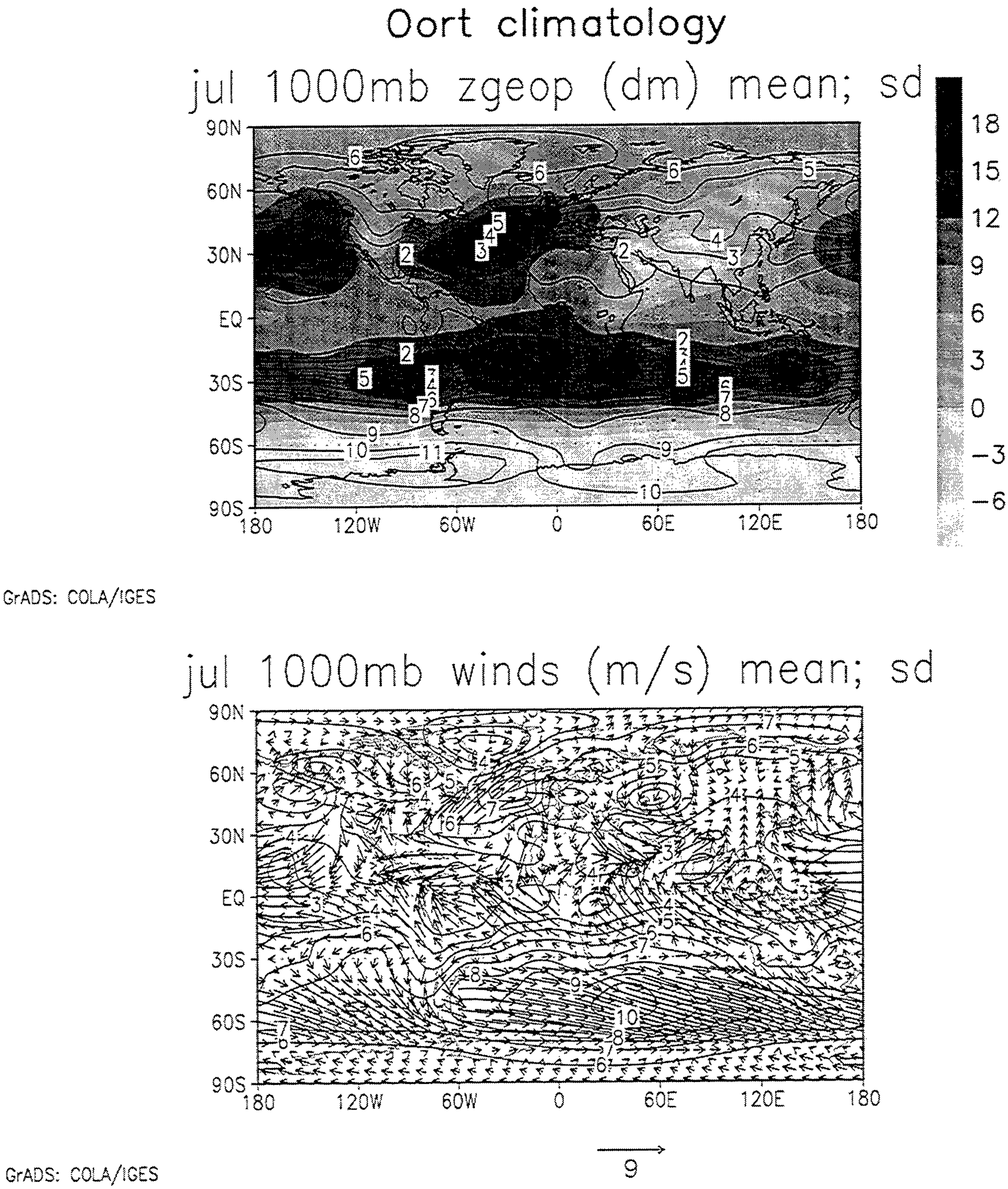


Figure 9: Oort climatology maps for July: 1000 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

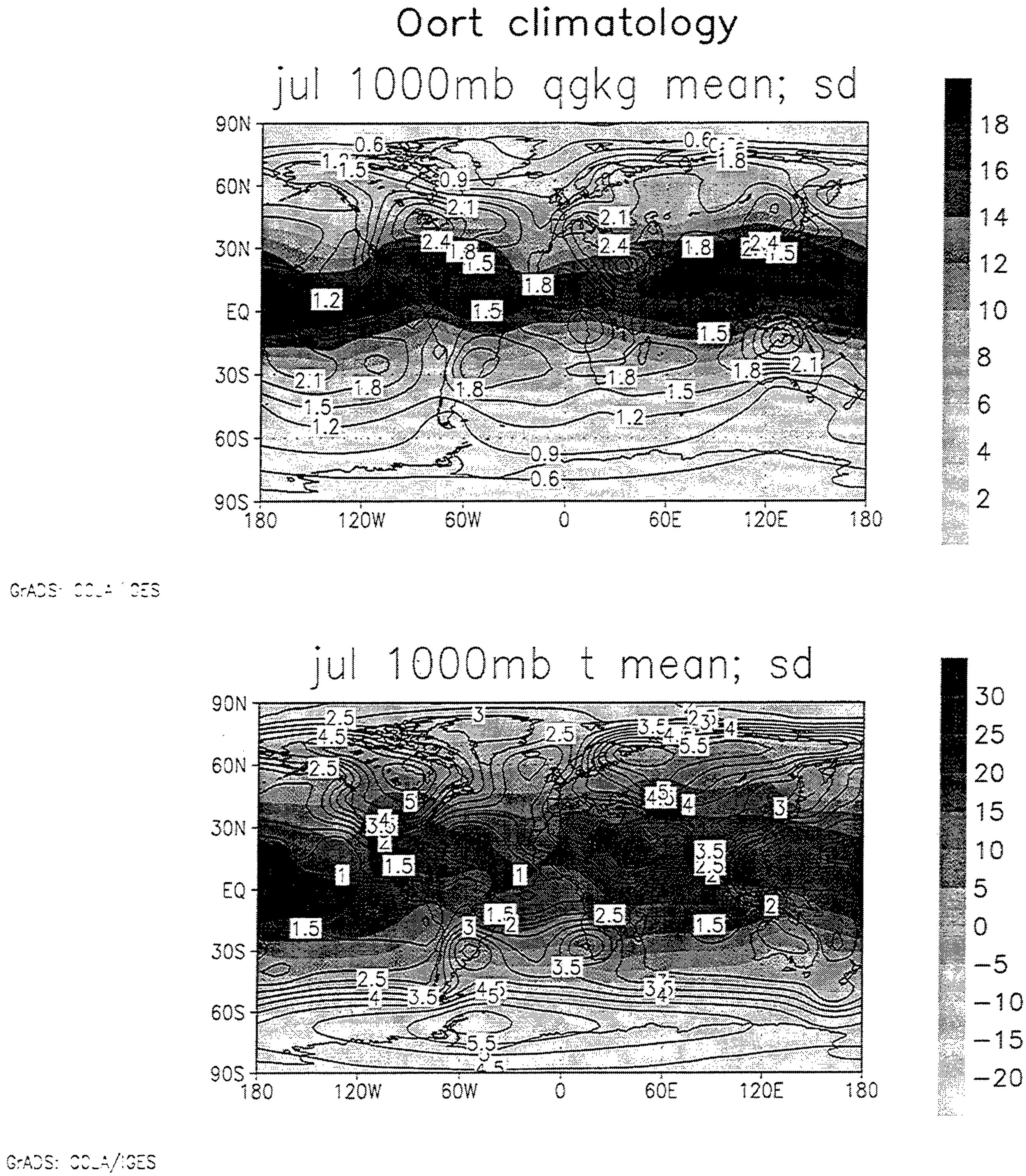


Figure 10: Oort climatology maps for July: 1000 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

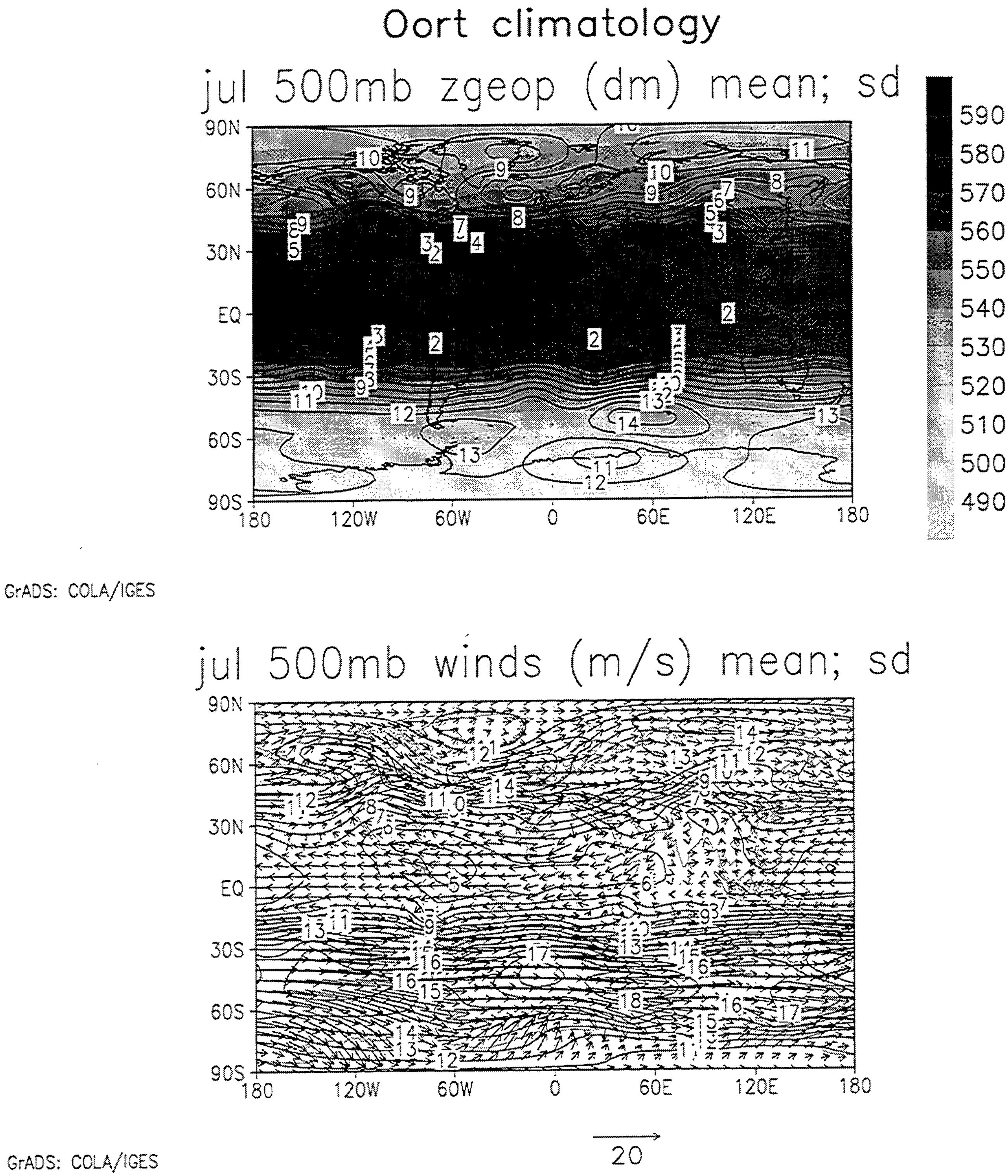


Figure 11: Oort climatology maps for July: 500 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

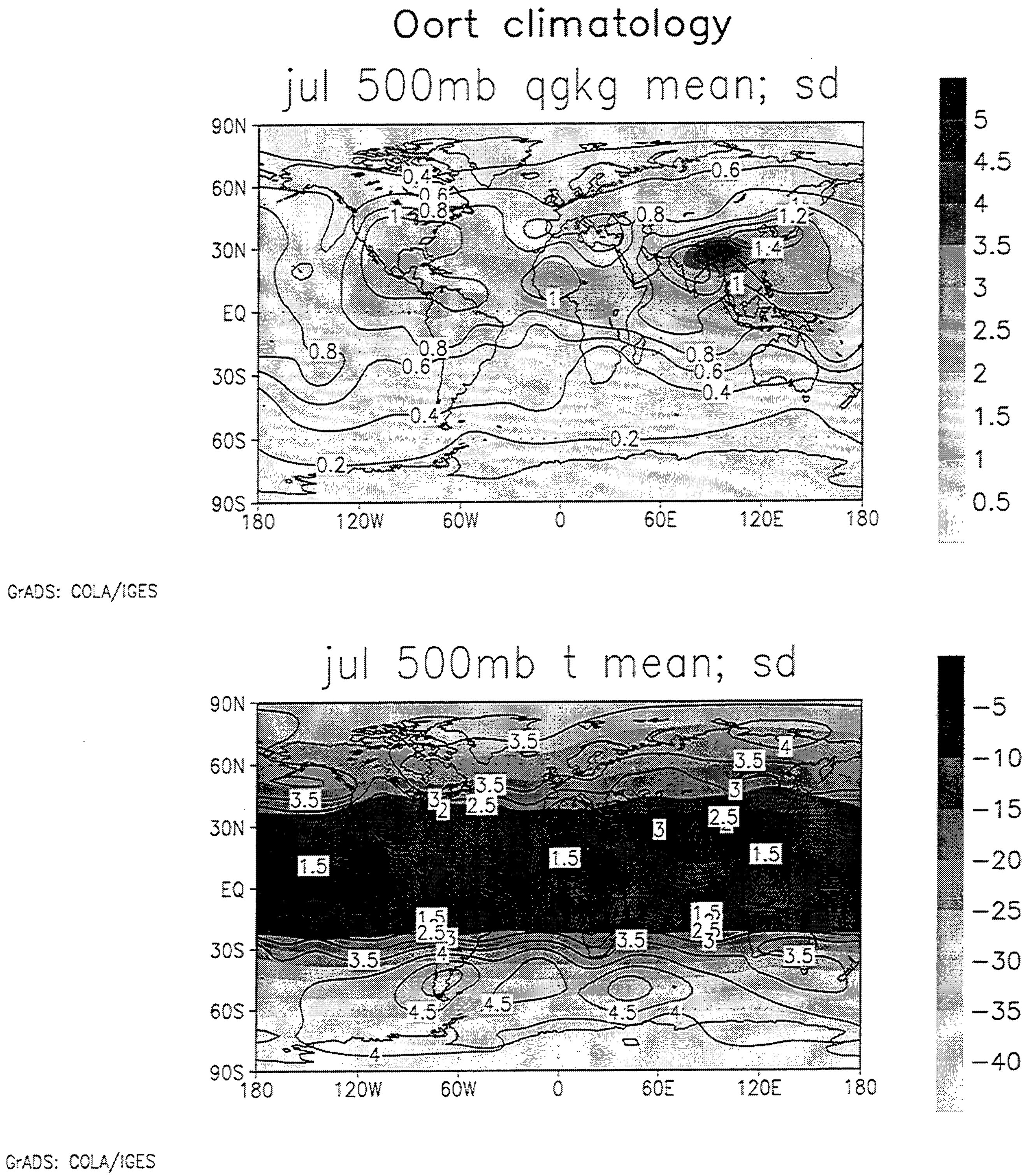


Figure 12: Oort climatology maps for July: 500 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

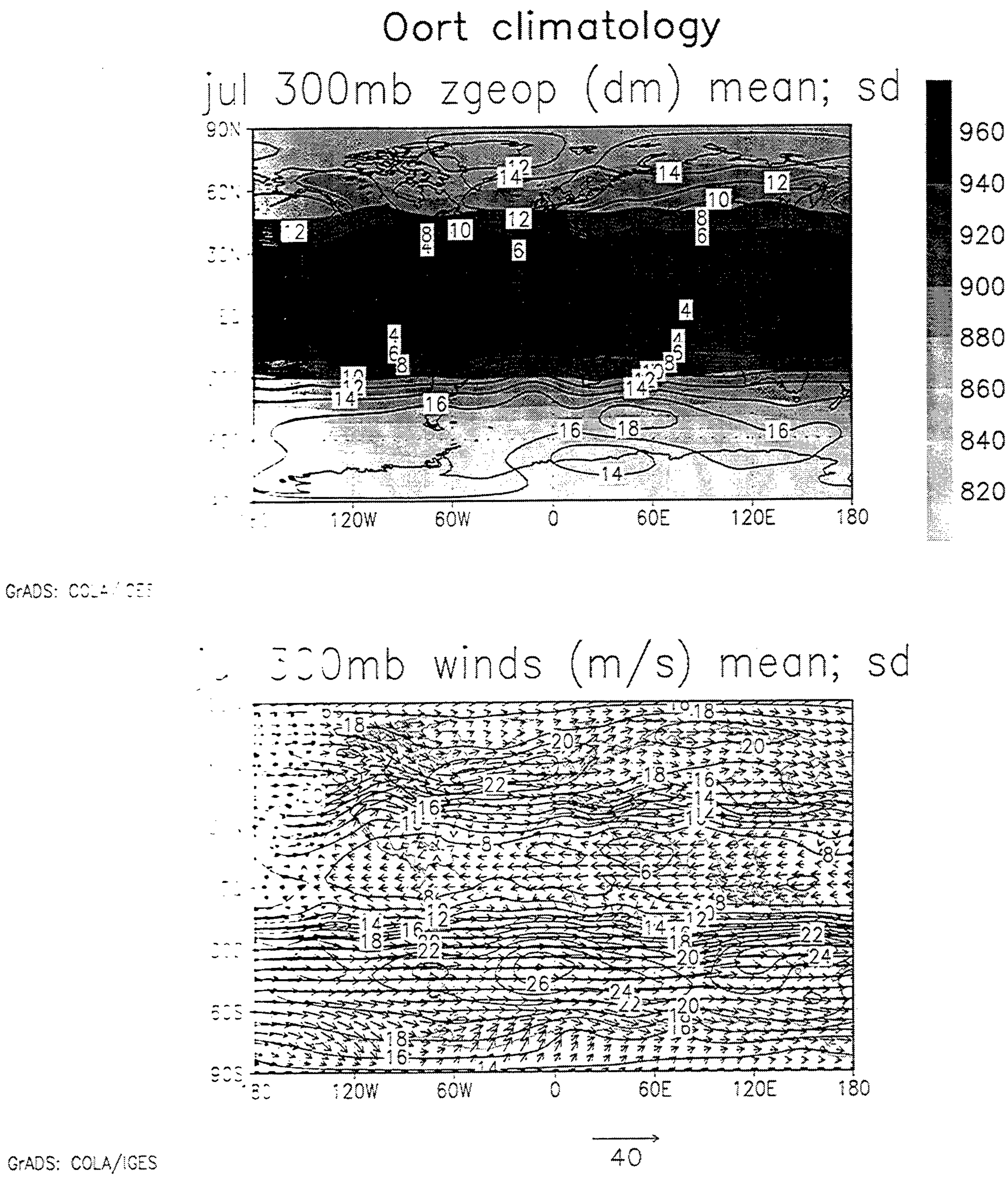


Figure 13: Oort climatology maps for July: 300 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

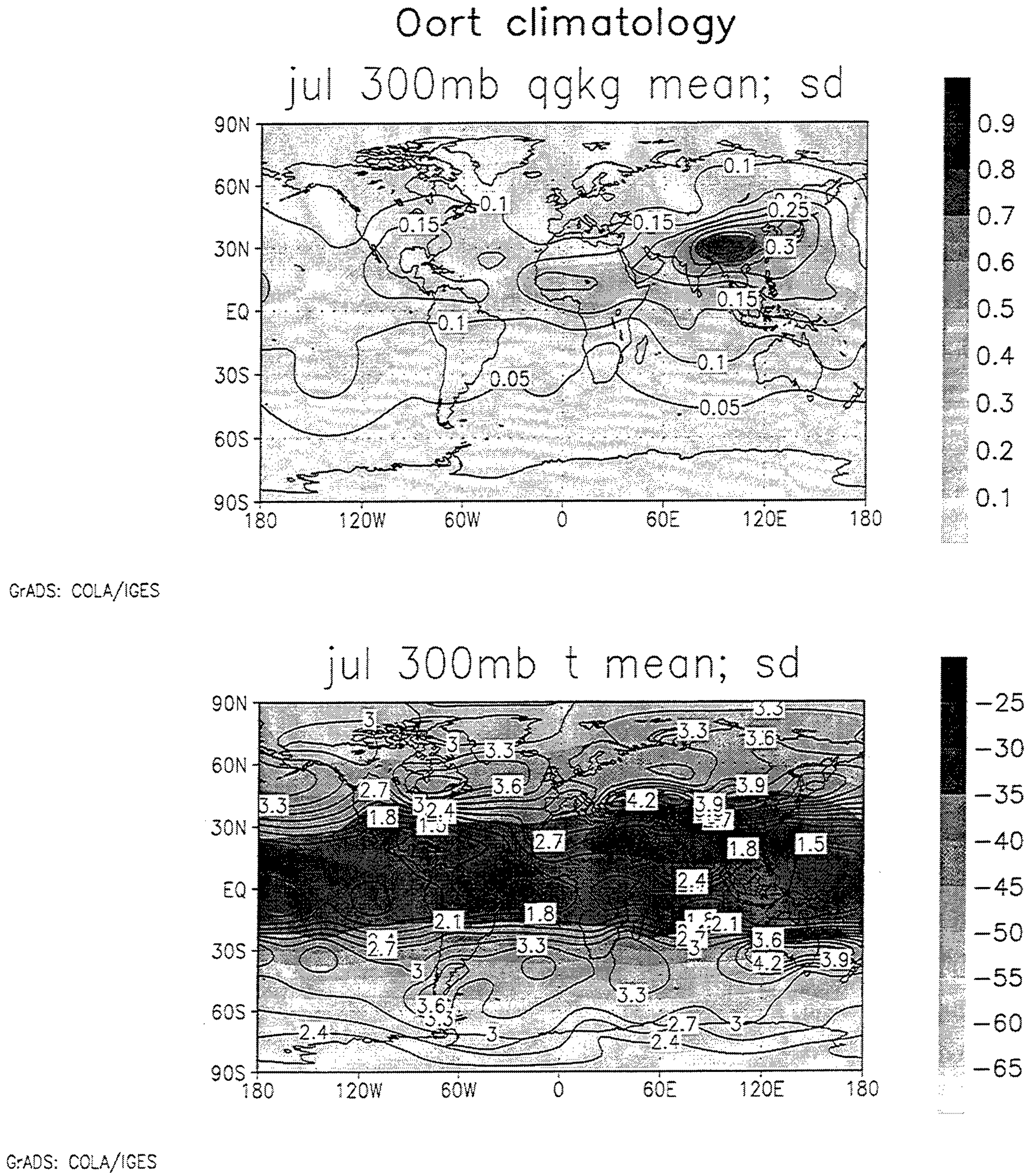


Figure 14: Oort climatology maps for July: 300 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

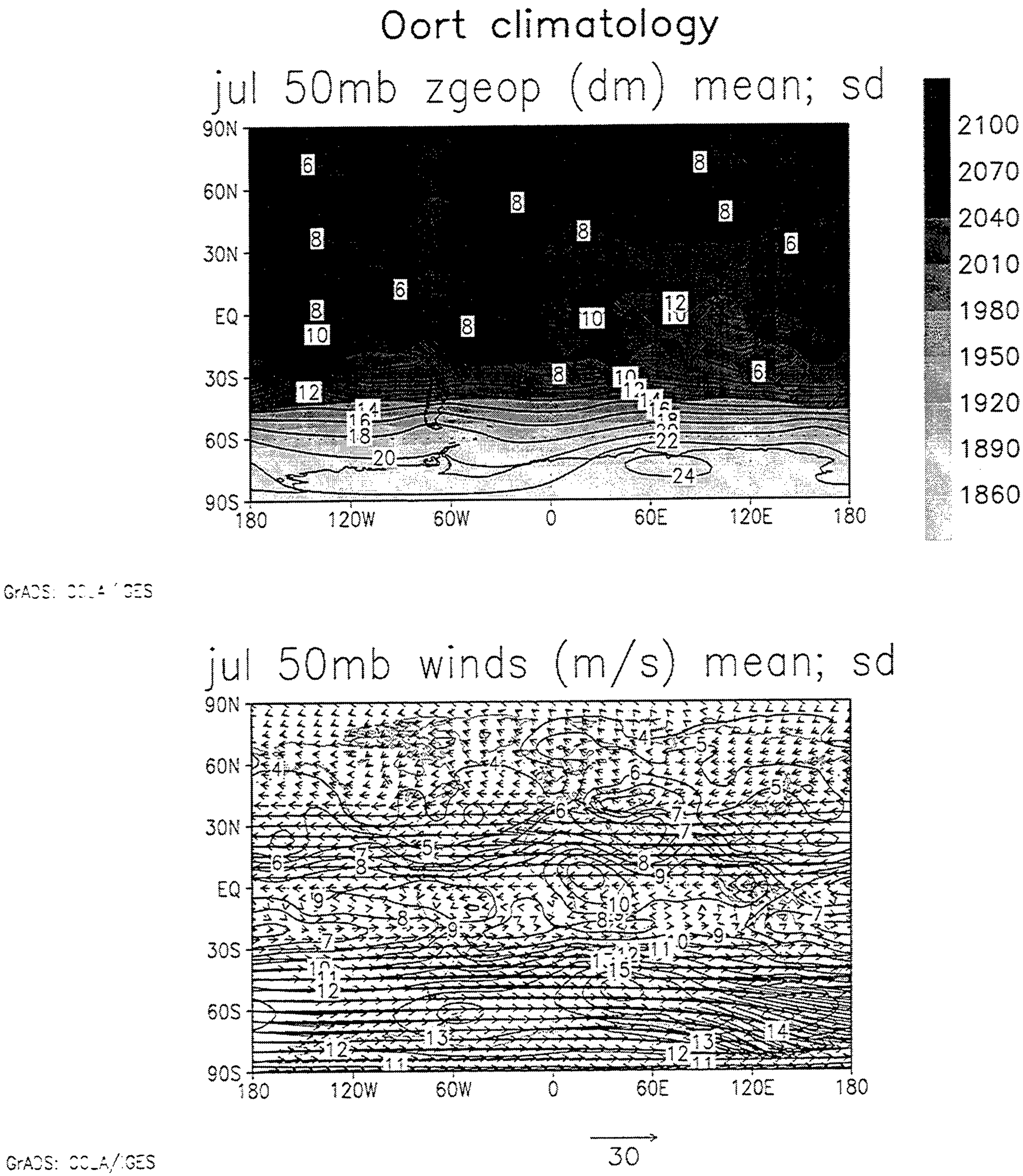


Figure 15: Oort climatology maps for July: 50 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

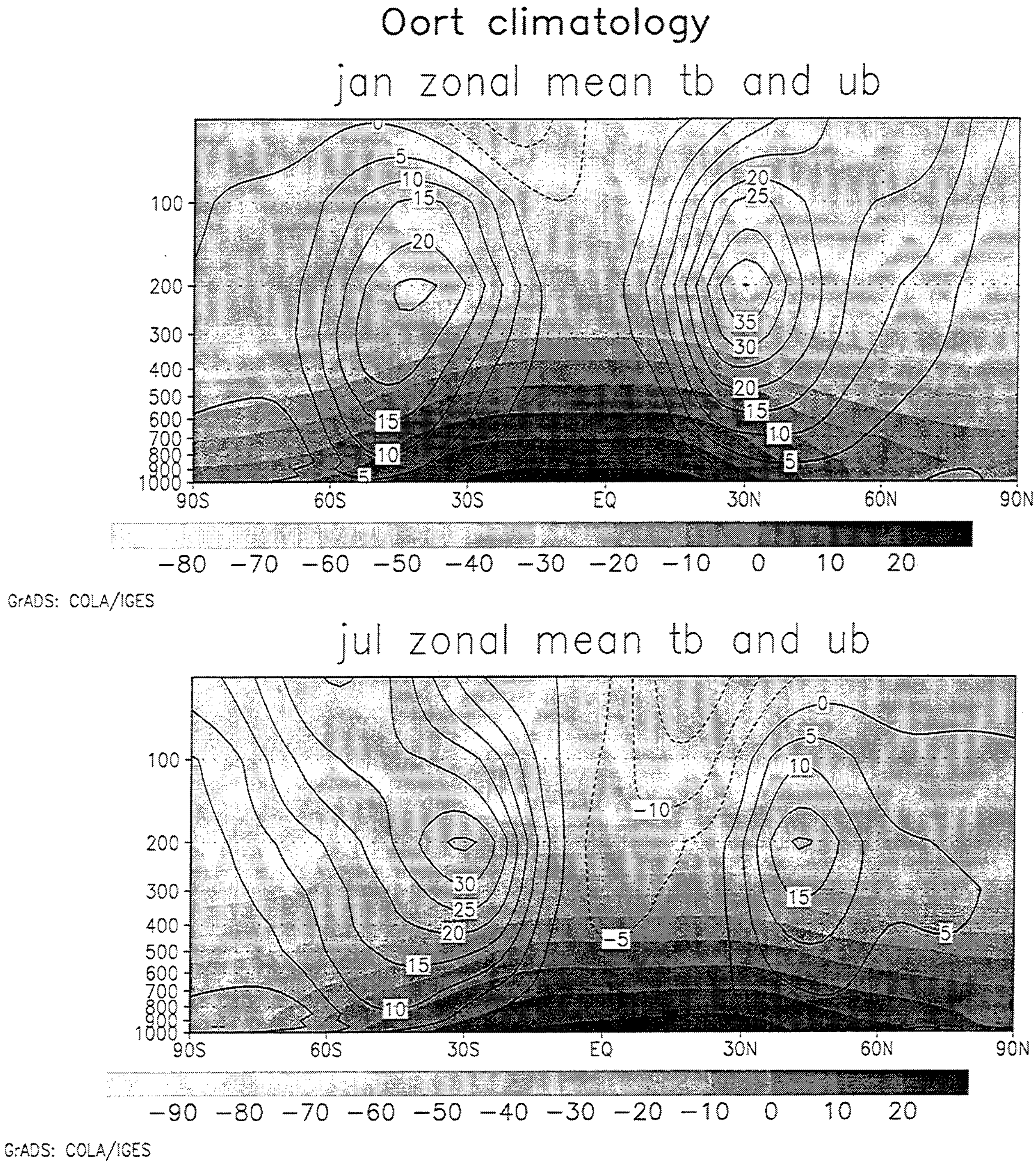


Figure 16: Oort climatology meridional cross sections. Zonal mean temperature (shaded) and zonal wind (contours) for January (top panel) and July (bottom panel).

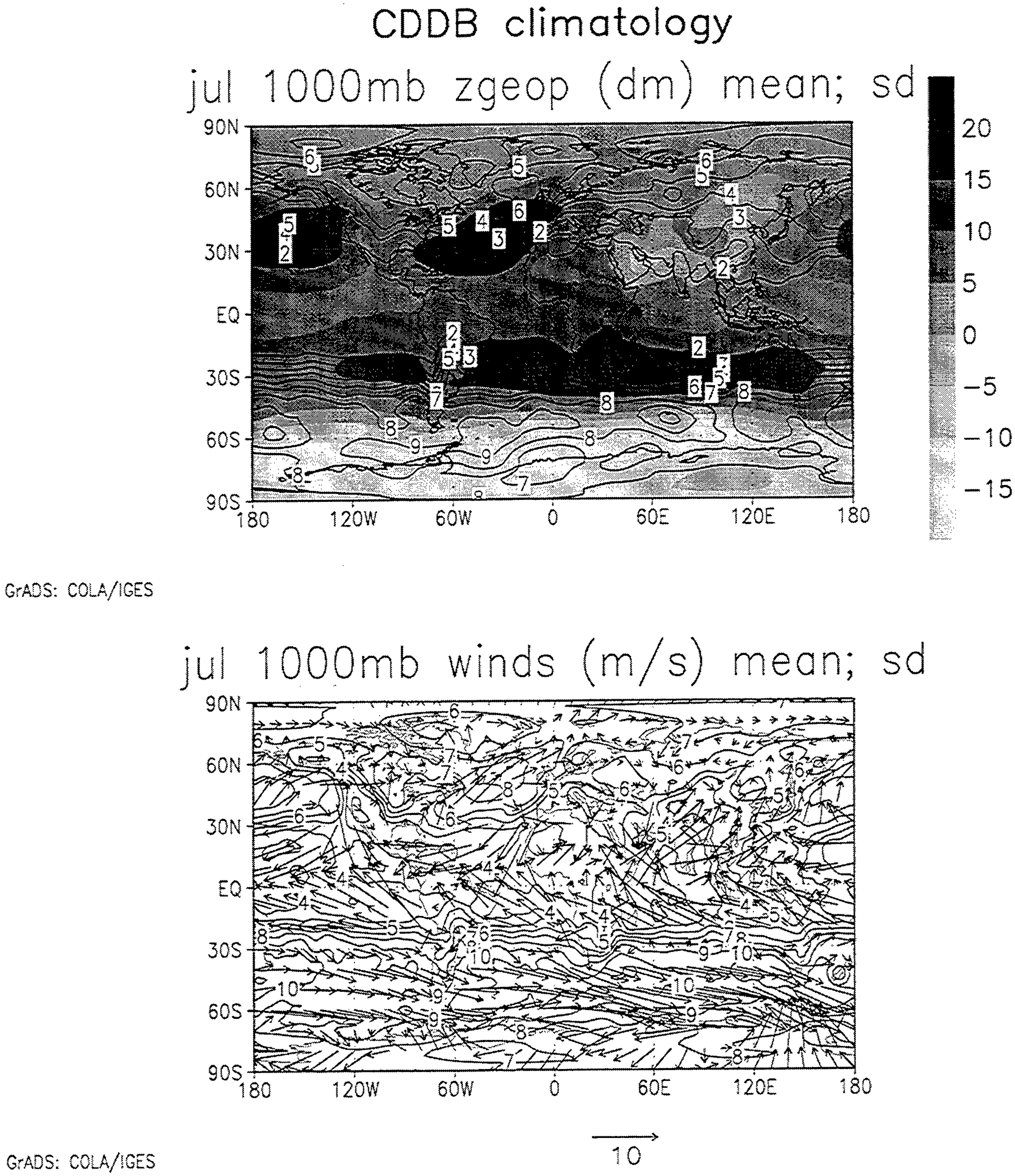


Figure 17: CDDB climatology maps for July: 1000 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

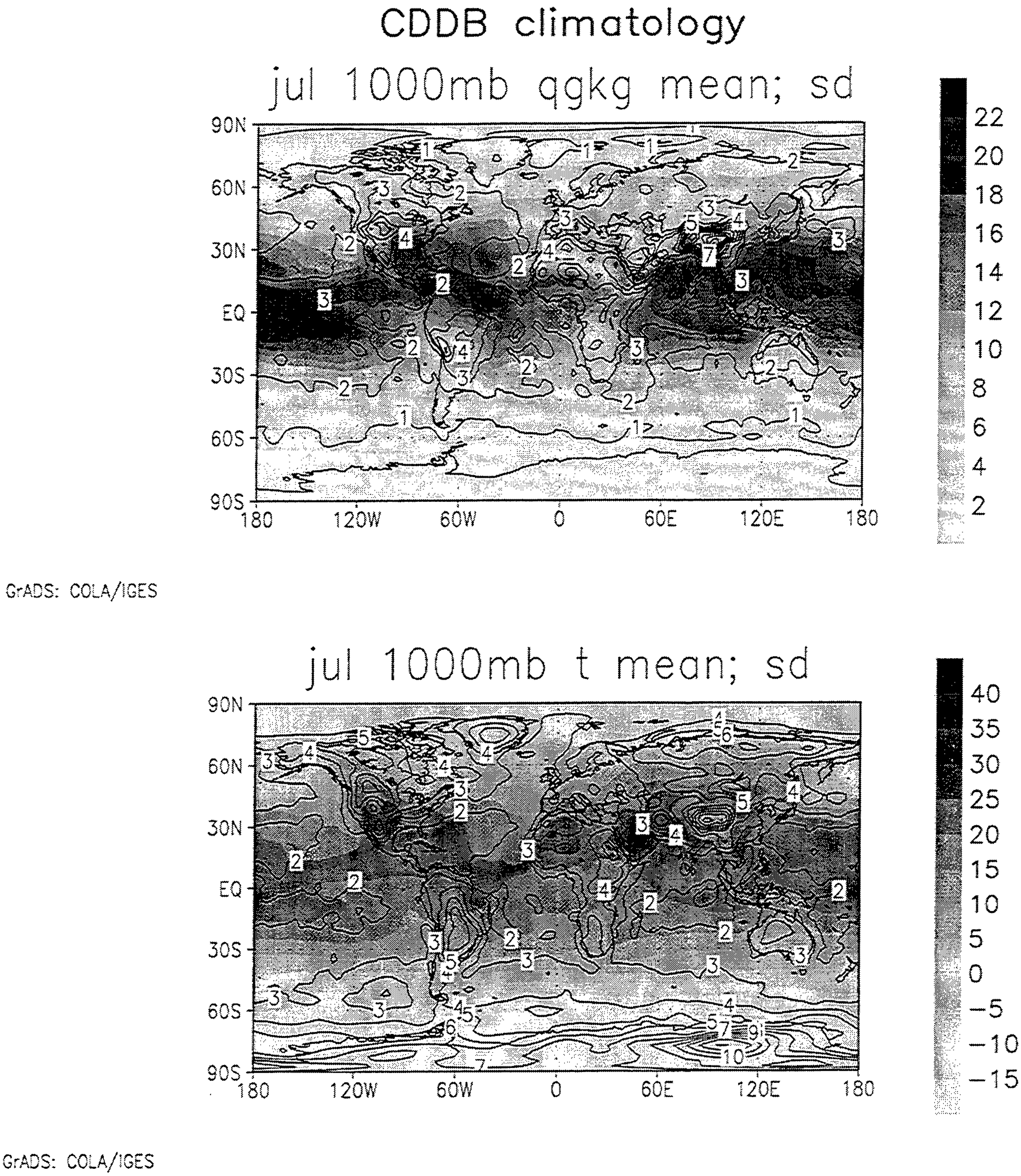
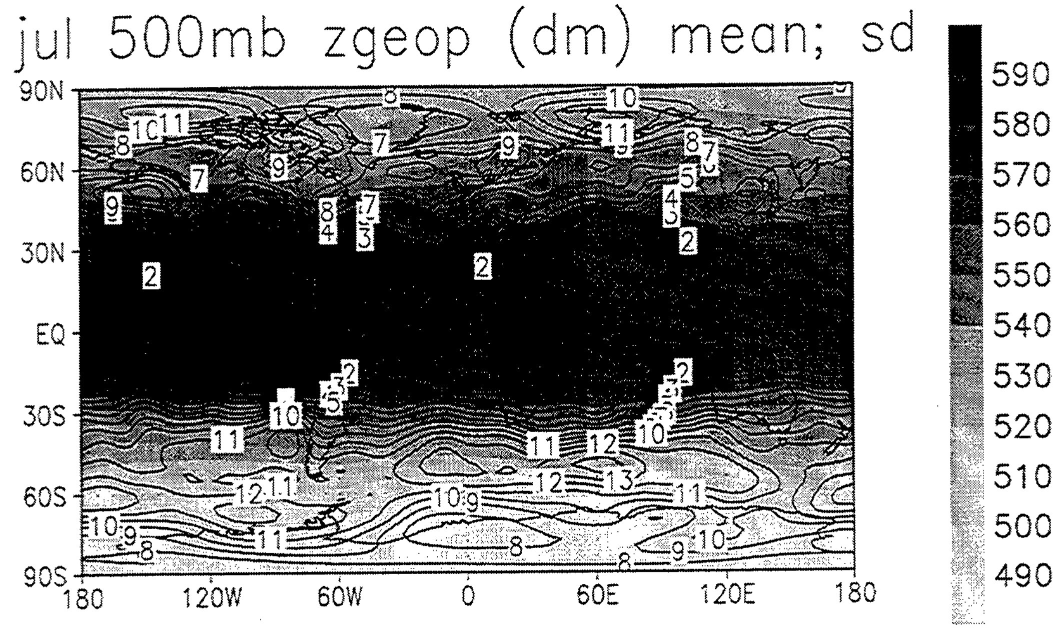
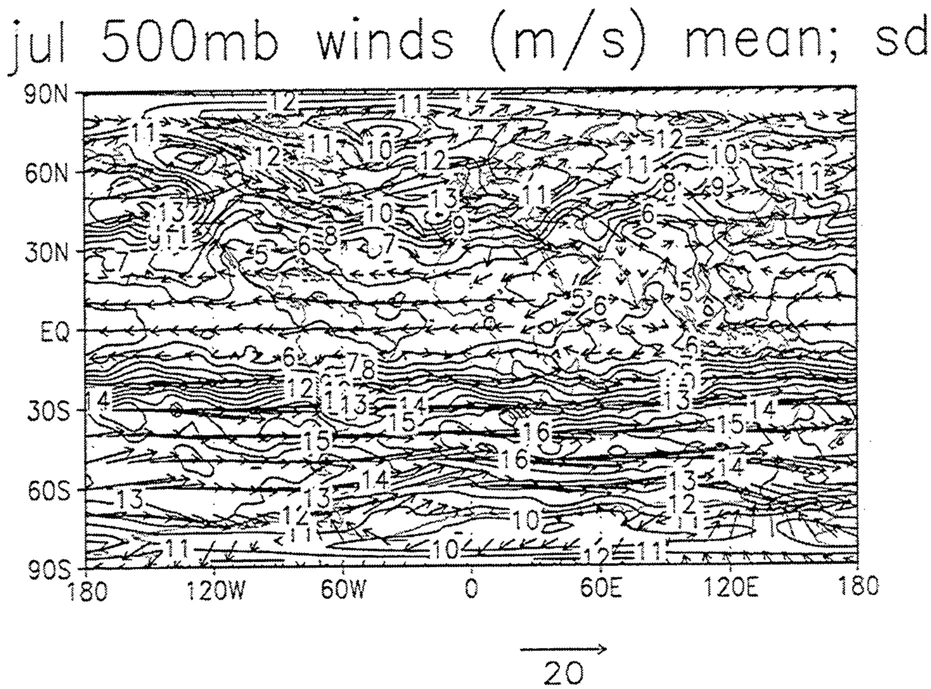


Figure 18: CDDB climatology maps for July: 1000 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

CDDB climatology



GrADS: COLA/IGES



GrADS: COLA/IGES

Figure 19: CDDB climatology maps for July: 500 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

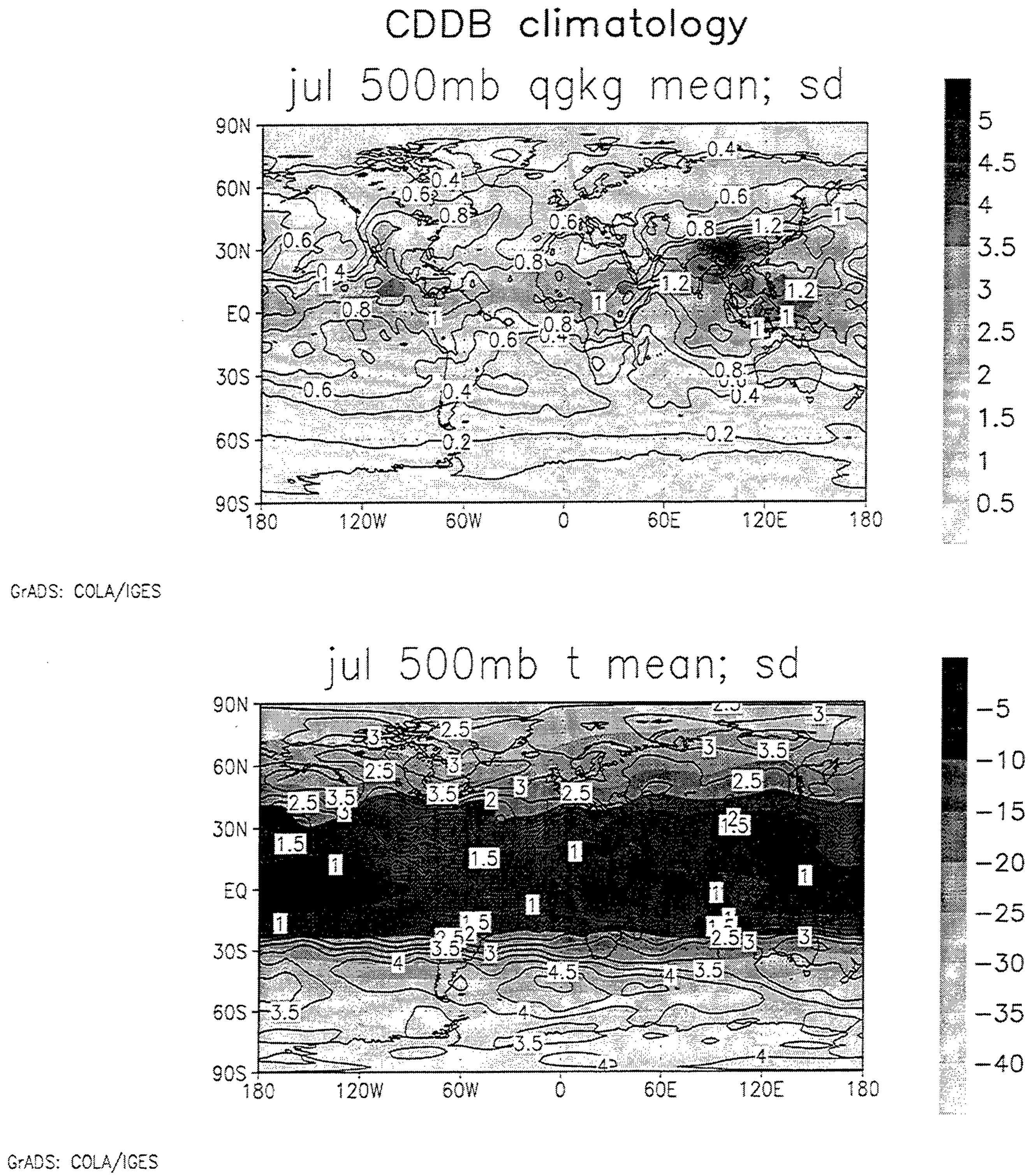
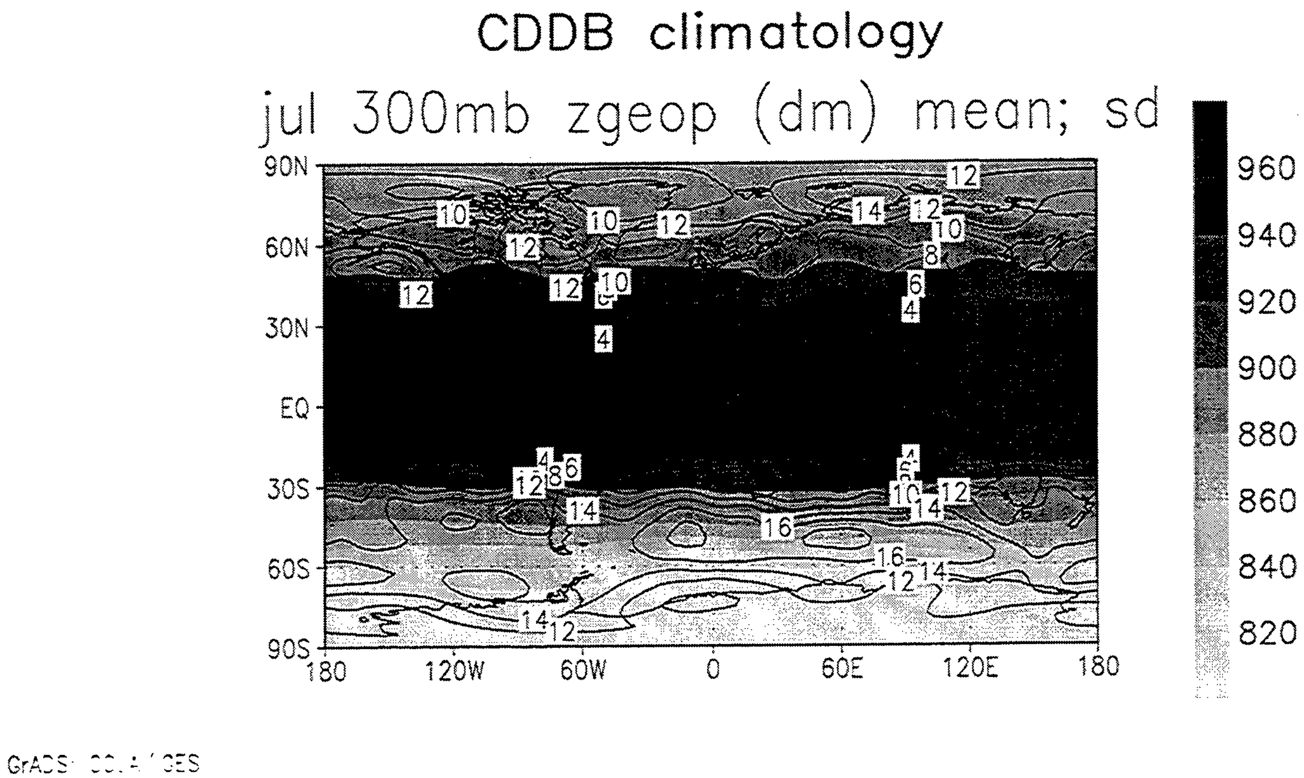


Figure 20: CDDB climatology maps for July: 500 hPa specific humidity and temperature. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).



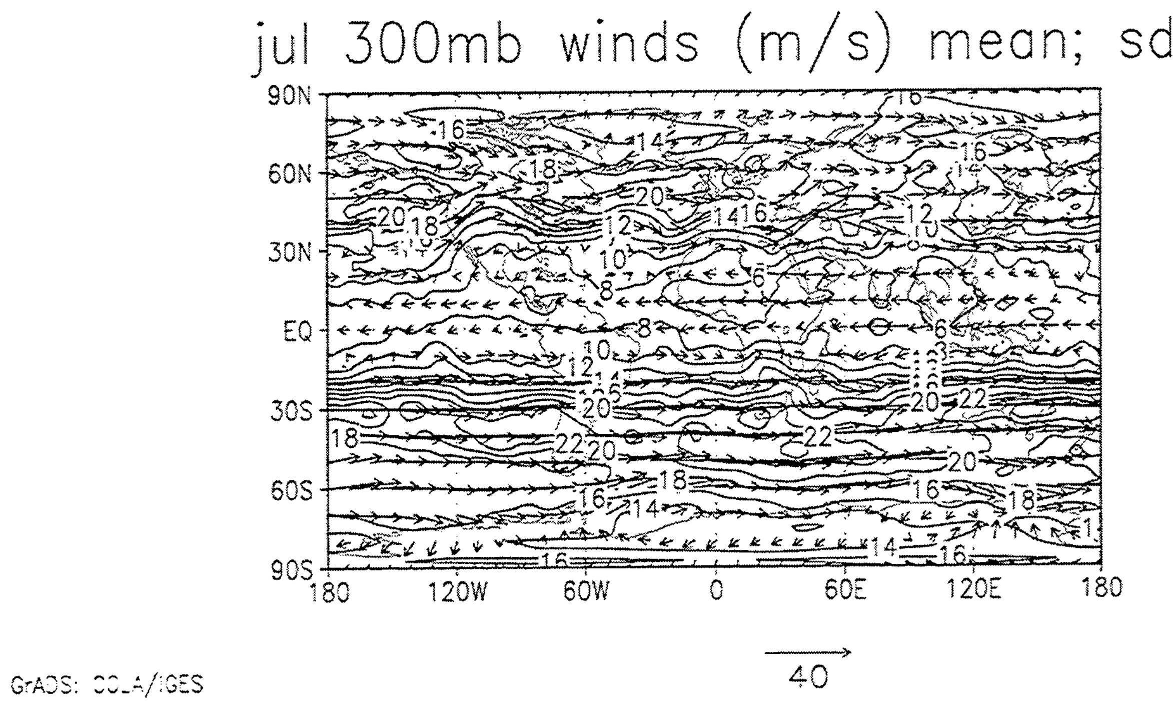


Figure 21: CDDB climatology maps for July: 300 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

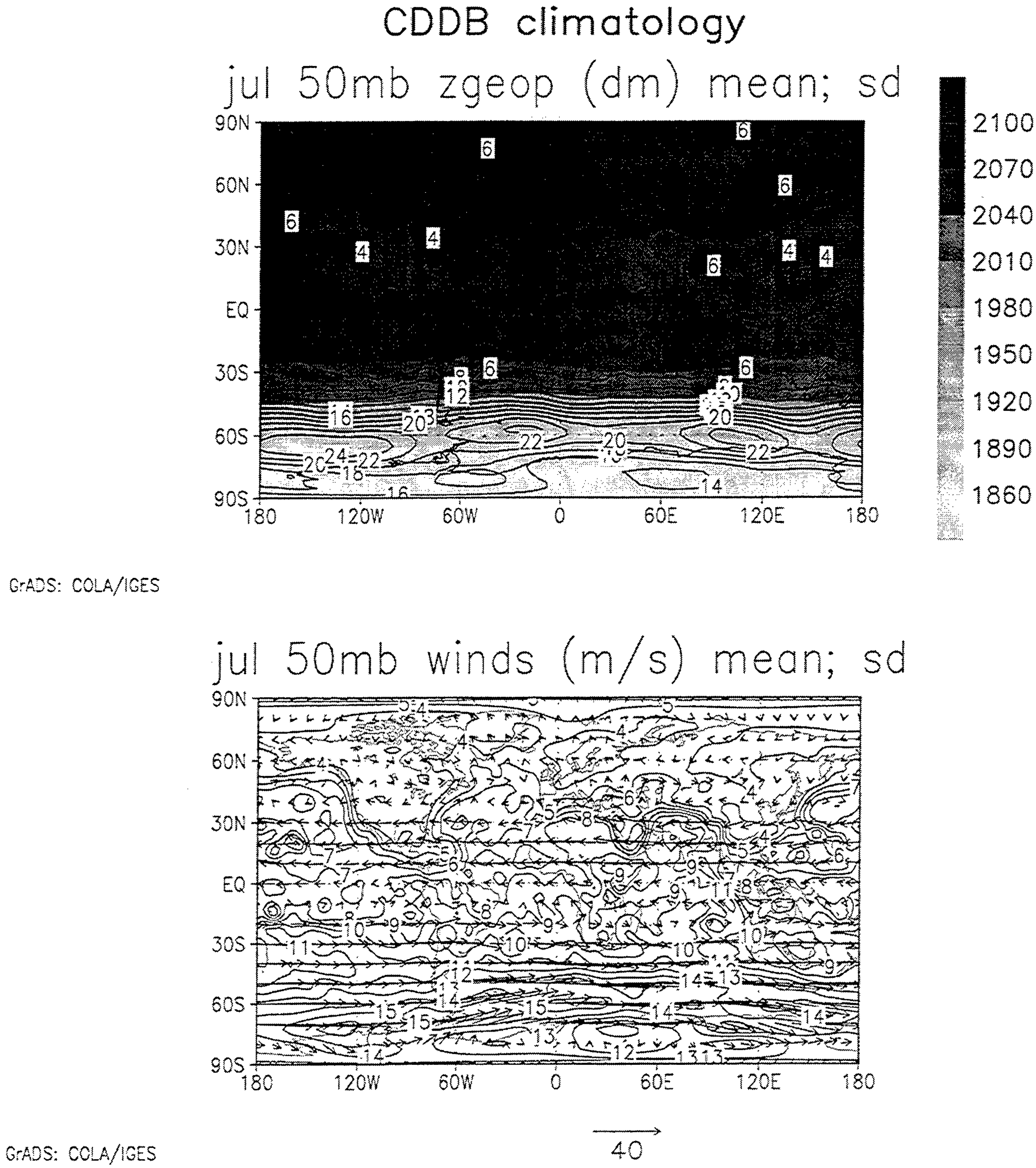


Figure 22: CDDB climatology maps for July: 50 hPa height and winds. Mean and standard deviations of geopotential (top panel) and vector winds (bottom panel).

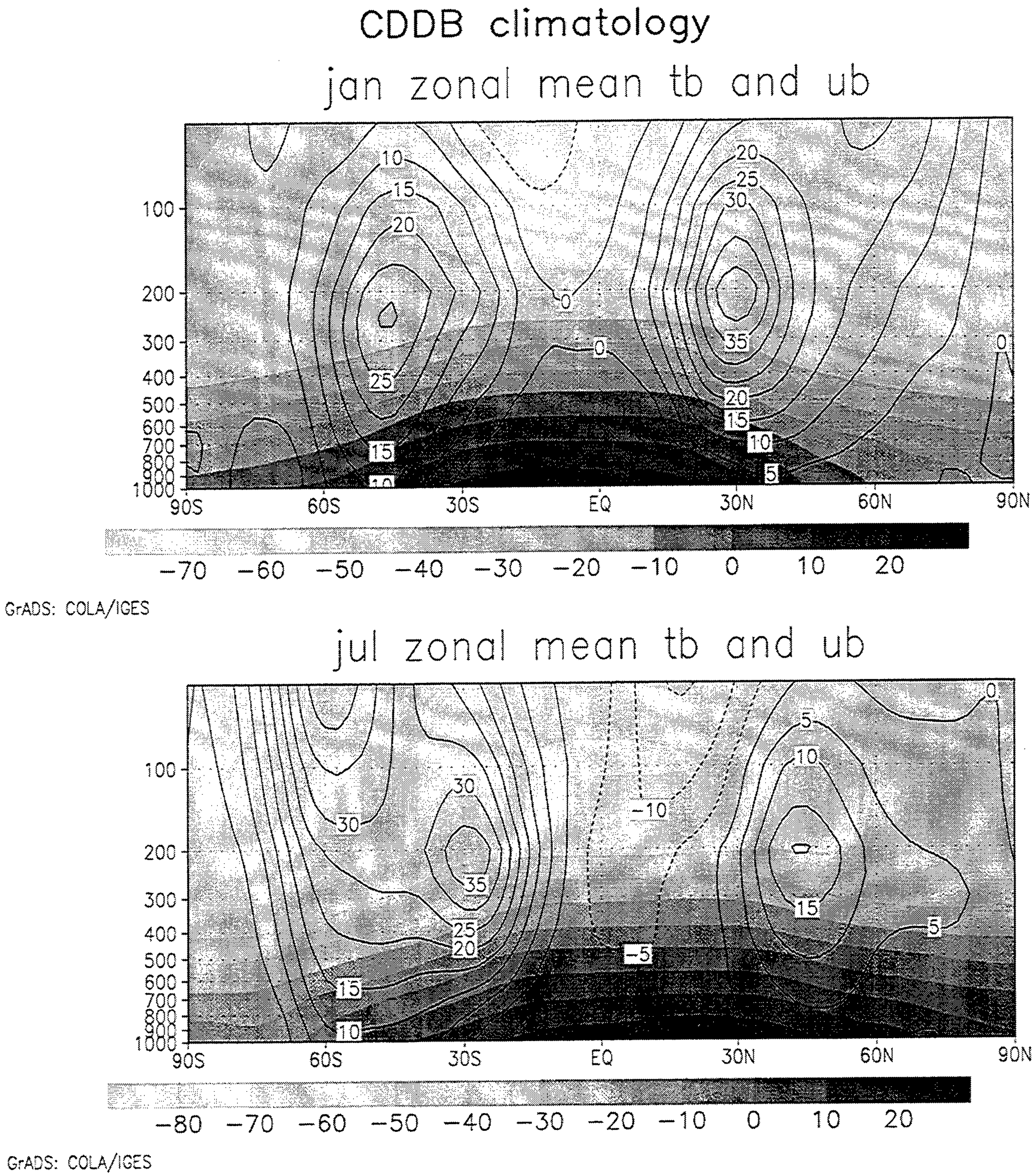


Figure 23: CDDB climatology meridional cross sections. Zonal mean temperature (shaded) and zonal wind (contours) for January (top panel) and July (bottom panel).

4 Error Statistics

Inherent in the modular and extensible design of TAP is the separation of the OI algorithm from the underlying observation and error statistics. The latter are specified through tables that can be tailored to the specific needs of the TAP application. An important part of the TAP project is the development of a statistical data base from which appropriate tables can be selected for different backgrounds, geographic areas, and observing systems.

Correlations of both forecast and observational errors are assumed to be separable and horizontally homogeneous. The error covariances are decomposed in the standard OI fashion:

$$C_{xy} = S_x S_y M_{xy} N_{xy}$$

where x and y are observed or forecast variables at observation or analysis locations; C_{xy} is the covariance between the errors of x and y; S_x is the standard deviation associated with x; S_y is the standard deviation associated with y; M_{xy} is the horizontal correlation between the errors of x and y; and N_{xy} is the vertical correlation between the errors of x and y. In practice, for global analyses the horizontal and vertical correlations may be allowed to vary slowly with vertical and horizontal location. In TAP, they vary geographically as well, but are fixed for a particular case, as described here: the horizontal correlations M_{xy} depend on distance and mean pressure and vertical correlations N_{xy} depend on the two pressures. To be precise, M_{xy} is represented by a set of 1–way tables in terms of horizontal distance, for different mean pressure levels, and N_{xy} is a set of 2–way tables in terms of the two pressures. In any particular case, these tables are interpolated as needed. The tables themselves are external to the analysis procedures. Observations of different type are assumed uncorrelated. Different tables are specified for each different observation and background type. For any analysis domain, appropriate versions of M_{xy} and N_{xy} are used. For some backgrounds and observation types, TAP includes different versions of the correlation tables, appropriate for the tropics and extratropics or for the continental and maritime situations. A single one of these is chosen when the analysis segment starts. Estimated prediction error standard deviations are also assumed to be constant in the horizontal in the OI development, but are often specified as a function of position. In TAP, background error standard deviations are either obtained as a gridded field (in the case of a climatology background), or specified for different geographical regions. In the latter case, a single set of standard deviations is chosen for the analysis domain.

During the first year of the TAP project, we conducted an extensive literature survey and assembled a set of appropriate references for the statistics of forecast model errors, observation errors, and deviations from climatology. An inventory of all references used in our literature survey was constructed, containing a description of the type of statistical information available from each reference. This inventory has been updated, and the updated information from this inventory is presented in the following section, organized by the type of statistics, with a brief description of the most relevant references for each type of statistic. Following this overview of error statistics references, the updated statistics selected for the baseline version of TAP are described in more detail.

4.1 Overview of Relevant Error Statistics References

4.1.1 Forecast Model Errors

4.1.1.1 Standard Deviations

[AndGM+86]: Figure of HIRLAM 12-hour 500 hPa height errors.

[BarM92]: Figure of Canadian Meteorological Centre operational model errors of heights and winds as a function of pressure.

[Ben89]: tables of background error standard deviations of height, temperature, relative humidity, and wind used in the MAPS (now called the Rapid Update Cycle, or RUC) isentropic analysis system. The background is NMC's 12–hour NGM forecast.

[BenSM+91, Car91, DevS94]: provide updated values for the statistics in [Ben89] for 3–hour forecasts from the RUC forecast model, for Montgomery potential, pressure, winds and humidity (condensation pressure).

[Ber79, McPBK+79]: NMC global prediction model errors for temperature, winds, and specific humidity at mandatory pressure levels. Prediction error growth rates for NMC global model

[DeyM85]: NMC global spectral model 6–hour forecast errors for temperature and winds at 12 mandatory pressure levels, for the extratropics and tropics.

[GoeP93]: NOGAPS global 6-hour forecast errors used in the Navy OI for height, wind, and thickness

[HolL86, LonH86]: ECMWF global grid point model 6-hour forecast errors for winds, height, and virtual temperature.

[Lor81, LonSU92]: ECMWF forecast errors used in the ECMWF OI for heights, winds, humidity, and thickness, separately for extratropics and tropics.

[MeuRA+90]: Values for the HIRLAM mesoscale forecast model errors for surface temperature and relative humidity.

4.1.1.2 Horizontal Correlations

[AndGM+86]: Functional fits of HIRLAM forecast error correlations for height, MSL pressure, temperature, and relative humidity. Anisotropy included for land/sea contrast.

[Ben89]: Second-order autocorrelation function fits of correlation on isentropic surfaces for NMC's 12–hour NGM forecast.

[Car91, SchC91, DevS94]: provide updated values for the correlations on isentropic surfaces for 3–hour forecasts from the RUC forecast model, for Montgomery potential, pressure, winds, and humidity (condensation pressure).

[DeyM85]: Functional fits for height and humidity error correlations.

[DiM88]: Functional fits for NMC global model error correlations, as used in the NMC RAFS.

[BarM92]: Figure of Canadian Meteorological Centre operational forecast model error correlations for 250 mb winds and 700 mb heights. Sample 2-d and 3-d correlation plots for heights, winds, height-winds.

[GoeP93]: NOGAPS global 6-hour forecast error correlations for heights, transverse and longitudinal winds, and height-wind cross-correlations.

[HolL86, LonH86]: ECMWF global grid point model 6-hour forecast error correlations for longitudinal and transverse winds, height, and wind-height cross-correlations.

[Lor81, LonSU92]: ECMWF forecast error correlations used in the ECMWF OI for heights, longitudinal and transverse winds, height-wind cross-correlations, humidity, and thickness, separately for extratropics and tropics.

[MeuRA+90]: Functional fits for the HIRLAM mesoscale forecast model error correlations for surface temperature, relative humidity, and winds.

[RogDD95]: Functional fits for the NMC Eta model error correlations for height and winds.

4.1.1.3 Vertical Correlations

[AndGM+86]: Table of HIRLAM forecast error correlations for height.

[Ben89]: Vertical correlation matrix for u-component of wind for NMC's 12–hour NGM forecast.

[DeyM85]: Functional fits for height error correlations.

[DiM88]: Functional fits for NMC global model error correlations, as used in the NMC RAFS.

[BarM92]: Canadian Meteorological Centre operational forecast model vertical error correlations for height, wind, and height-wind correlations.

[HolL86, LonH86]: ECMWF global grid point model 6-hour forecast errors vertical correlations for longitudinal and transverse winds, and height.

[LonSU92]: ECMWF forecast error correlations used in the ECMWF OI for heights and longitudinal and transverse winds.

[GoeP93]: NOGAPS global 6-hour forecast error vertical correlations for heights and transverse and longitudinal winds.

[RogDD95]: Functional fits for the NMC Eta model error correlations for height and winds.

4.1.2 Climatology Background Errors

4.1.2.1 Standard Deviations Background error standard deviations for the climatology background are obtained from the climatological variance itself, which is available in gridded form (see Nehrkorn *et al.* 1995 [NehHY95]).

4.1.2.2 Horizontal Correlations

[AndGM+86]: Plots and functional fits of MSL pressure and surface temperature correlations

[Bue71, Bue72]: Plots and functional fits of 500 mb and 200 mb wind and height correlation for summer and winter over Northern Hemisphere continents.

[JulT75, Thi75, Thi76, Thi77, Thi85]: 500 mb height, temperature, and wind correlations and functional fits.

[RamKS73]: 500 mb wind correlations over Indian region.

4.1.2.3 Vertical Correlations No references have been found with tables or functional fits for vertical correlations of climatology background errors. Computation of these statistics is possible from archived datasets, but has been postponed for later phases of the project.

4.1.3 Observation Errors

4.1.3.1 Standard Deviations

[AndGM+86]: Tables of OESDs for height, temperature, relative humidity, and wind for rawinsonde, aircraft winds, SATEM thicknesses, surface pressure and winds, and satellite winds.

[Ben89]: OESDs for height, temperature, relative humidity, and wind for rawinsonde, profiler, aircraft, and surface observations.

[Ber79]: Temperature and wind errors for rawinsonde, aircraft, satellite retrievals, and cloud drift winds.

[Car91]: rawinsonde OESDs for Montgomery potential, pressure, winds, and humidity (condensation pressure) at several isentropic levels.

[DeyM85]: Temperature and wind errors for rawinsonde, aircraft, satellite retrievals, and cloud drift winds.

[GoeP93]: Temperature, height, thickness, and wind errors for rawinsonde, aircraft, satellite retrievals, cloud drift winds, and surface observations.

[LonSU92]: Temperature, height, thickness, wind, and humidity errors for rawinsonde, aircraft, satellite retrievals, and cloud drift winds, surface observations, drifting buoys, and pilot balloons.

[Lor81]: Temperature, height, thickness, and wind errors for rawinsonde and aircraft observations.

4.1.3.2 Horizontal Correlations

[AndGM+86]: Functional fit for satellite derived thicknesses.

[DeyM85]: horizontal correlations of satellite retrieved thickness errors.

[LonSU92]: horizontal correlations of satellite retrieved thickness errors.

4.1.3.3 Vertical Correlations

[Ber79]: Rawinsonde error vertical correlations for winds and geopotential. Satellite height error correlations.

[HolL86, LonH86]: Rawinsonde error vertical correlations for winds, geopotential, and thickness

[LonSU92]: Rawinsonde error vertical correlations for geopotential, satellite retrieval thickness error vertical correlation.

[GoeP93]: Radiosonde height error vertical correlation.

4.2 TAP Baseline Error Statistics References

The error statistics for the baseline version of TAP rely most heavily on the statistics used in NOGAPS (Goerss and Phoebus, 1993 [GoeP93]). The primary reason is that, at the beginning of this project, the most likely scenario for a TAP implementation was for global forecast model background fields obtained from the Navy NOGAPS model. The NOGAPS model error statistics were thus most relevant to TAP. Furthermore, the Navy NOGAPS system is a state-of-the-art operational system with recent and fairly complete and accessible documentation. These statistics have been supplemented where needed with information from other sources from the above list, in most cases from the ECMWF documentation found in Lönnberg *et al.* (1992 [LonSU92]). Elements for which no source has yet been selected have been marked by "TBD". Where no entry for correlations exists, autocorrelations are assumed to be zero except at zero separation, and crosscorrelations are assumed to be zero.

4.2.1 Forecast Model Errors

4.2.1.1 Standard Deviations These standard deviations have all been derived from comparisons with radiosonde data over well-sampled regions. As a possible enhancement, these values could be inflated by an appropriate error growth rate over data-sparse regions.

height: [GoeP93, Tables on p. 41]

temperature: Values computed from height error covariances

thickness: computed from height error covariances

winds: [GoeP93, Tables on p. 41]

relative humidity: [LonSU92, p.3.5]

4.2.1.2 Horizontal Correlations

height: [GoeP93, eq. 30, p.12-13]

temperature: use values for height from [GoeP93, eq. 30, p.12-13]

thickness: computed from height error covariances

winds: compute from height correlations, as in [GoeP93, eq. 31-41, p.13-14]

height–winds crosscorrelations: compute from height correlations, as in [GoeP93, eq. 31-41, p.13-14]

relative humidity: [LonSU92, p. 3.5]

4.2.1.3 Vertical Correlations

height: [GoeP93, p.18]

thickness: computed from height error covariances

temperature: computed from height error covariances

winds: use vertical correlation function for height

4.2.2 Climatology Model Errors

4.2.2.1 Standard Deviations Background error standard deviations for the climatology background are obtained from the climatological variance itself, which is available in gridded form (see Nehrkorn *et al.* 1995 [NehHY95])

4.2.2.2 Horizontal Correlations

height: [JulT75, Tables 1 and 2, function R2.1]

temperature: use values for height from [JulT75]

thickness: computed from height error covariances

winds: compute from height correlations, as in [GoeP93, eq. 31-41, p.13-14]

height–winds crosscorrelations: compute from height correlations, as in [GoeP93, eq. 31-41, p.13-14]

relative humidity: use values from [LonSU92, p. 3.5]

4.2.2.3 Vertical Correlations

height: use the same functional form as [GoeP93, p.18] with an adjusted vertical length scale ($= 6500m$)

thickness: computed from height error covariances

temperature: computed from height error covariances

winds: use vertical correlation function for height

4.2.3 Observation Errors

The observation errors in these references have not been adjusted for the representational error (*i.e.* the representational error has not been separated from the instrument error). This is left as a future enhancement to the statistical database.

4.2.3.1 Standard Deviations

rawinsonde: **height:** [GoeP93, Tables on p.41]

- **temperature:** a nominal value of 0.5 K is used
- **winds:** [GoeP93, Tables on p.41]
- **relative humidity:** [LonSU92, p.3.1]

aircraft winds: [GoeP93, Tables on p.41]

satellite retrievals: **thickness:** [GoeP93, Tables on p.41]

- **temperature:** [LonSU92, p.2.36]
- **relative humidity:** [LonSU92, p.3.4]

cloud drift winds: [GoeP93, Tables on p.41]

surface observations: **height:** [GoeP93, Tables on p.41]

- **winds:** [GoeP93, Tables on p.41]
- **temperature:** a nominal value of 0.5 K is used
- **relative humidity:** [LonSU92, p.3.1-3.3]

4.2.3.2 Horizontal Correlations

satellite thickness: [LonSU92, p. 2.37]

cloud drift wind: TBD

4.2.3.3 Vertical Correlations

rawinsonde height: [LonSU92, p. 2.38]

satellite thickness: The values given in [LonSU92, p. 2.39] were negligibly small. The TAP baseline version uses zero vertical correlations.

4.3 TAP Baseline Error Statistics Data

In this section, we describe the error statistics models of the TAP baseline error statistics references, to the extent that they have been implemented in TAP.

4.3.1 Standard Deviations

Summary plots of the error standard deviations are shown in Figure 24 for height errors of the forecast first guess and Raob reports, and thickness errors of clear and cloudy satellite retrievals. Note that the SATEM thickness errors are plotted at the bottom of each layer—their vertical variation depends on the thickness of each layer and the accuracy of the retrieved layer-mean temperature. Figure 25 contains the temperature errors of the forecast and Raobs, Figure 26 the wind errors of the forecast, Raobs/Pibals, aircraft reports, and cloud-drift winds. Observation errors that do not depend on altitude are given in Table 1.

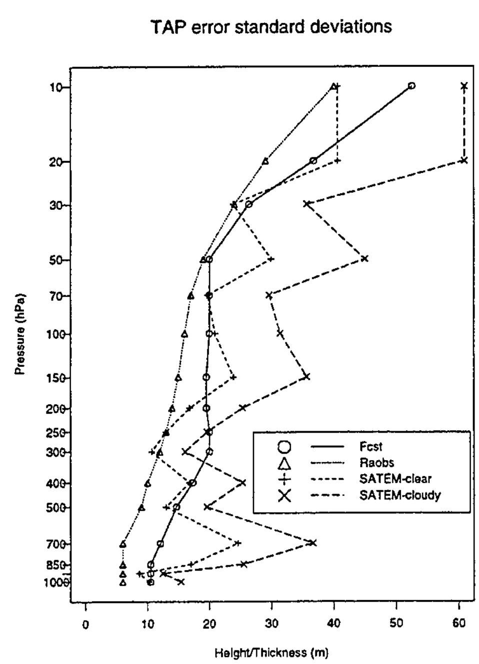


Figure 24: Height and thickness (for SATEMs) error standard deviations used in TAP.

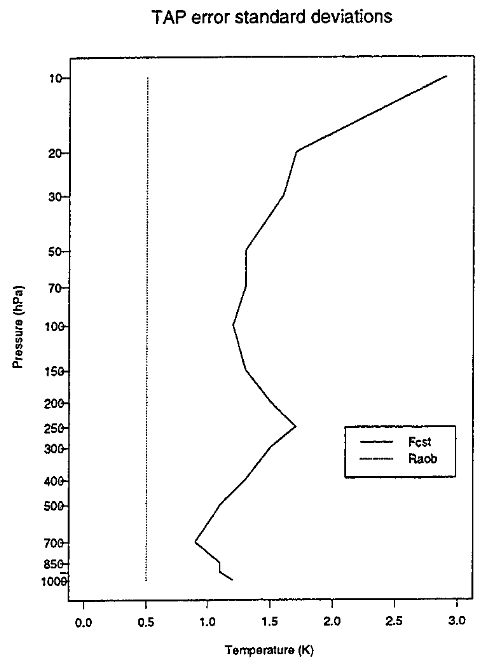


Figure 25: Temperature error standard deviations used in TAP.

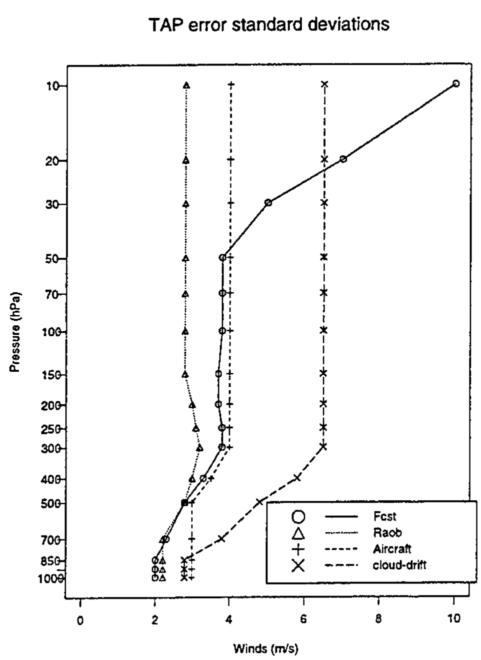


Figure 26: Wind error standard deviations used in TAP.

Table 1: Observation errors used in TAP. SATEM indicates satellite retrievals, Surface all surface land, ship, and buoy observations, and Paobs manually bogused observations of tropical storms based on satellite imagery.

Type	Height m	Winds m/s	Temperature K	Relative Humidity %
SATEM				15
Raob				15
Surface	6	2.2	0.5	
SSM/I		2.2		
Paobs	24.1			

4.3.2 Horizontal correlations

The functional fit of the height-height error correlations for the NOGAPS forecast model is given by the following modified second-order autoregressive (SOAR) function:

$$\rho_{zz}(r) = 1 - c_2(1 + c_1)exp(-c_1 r) \ ,$$

where r is the great circle separation distance, and c_1 and c_2 are constants ($c_1 = 2.6 \cdot 10^{-3} km^{-1}$ and $c_2 = 0.9$). The function is shown in Figure 27. The resulting contour plot of correlation values is shown in Figure 28 over a map background.

In the present version of TAP, the height-height correlation function is also used for the computation of mass-wind and wind-wind correlations. Only a brief outline of the procedure is given here – the mathematical details can be found in the Software Requirements Specification and Software Design Document, which are provided as separate contract deliverables. All correlations involving wind components are computed in terms of the natural coordinate system (longitudinal $\tilde{u}$ and transverse $\tilde{v}$) components, and converted to the local (east and north) coordinate system as needed. Using assumptions of isotropy and homogeneity, the wind correlations can then be related to those of streamfunction (ψ) and velocity potential (χ). TAP also employs the common assumption that $\rho_{\psi\psi} = \rho_{\chi\chi}$ and $\rho_{\psi\chi} = 0$. For the baseline set of error statistics, the horizontal structure functions of streamfunction and velocity function are not specified independently, but instead set equal to the height-height autocorrelation function. If we define a length scale L from

$$L^2 = \frac{-2\rho_{zz}(r=0)}{\nabla^2 \rho_z(r=0)} = \frac{-2\rho_{zz}(r=0)}{(\frac{1}{r}\frac{d}{dr} + \frac{d^2}{dr^2})\rho_{zz}(r=0)} \ ,$$

and rescaled first and second derivatives of the horizontal structure function as

$$\begin{aligned} f &= -\frac{L^2}{r}\frac{d}{dr}\rho_{zz}, \\ g &= -L^2\frac{d^2}{dr^2}\rho_{zz}, \end{aligned}$$

we can write the wind-wind correlations as follows:

$$\begin{aligned}\rho_{\tilde{u}\tilde{u}} &= (1-\nu^2)N_{\psi\psi}f + \nu^2 N_{\chi\chi}g,\\ \rho_{\tilde{v}\tilde{v}} &= (1-\nu^2)N_{\psi\psi}g + \nu^2 N_{\chi\chi}f,\\ \rho_{\tilde{v}\tilde{u}} &= \rho_{\tilde{v}\tilde{u}} = 0,\end{aligned}$$

where $N_{\psi\psi}$ and $N_{\chi\chi}$ are the vertical correlation functions for streamfunction and velocity potential, and the parameter ν^2 can be specified to control the partitioning between rotational and divergent kinetic energy ($0 \leq \nu^2 \leq 1$). As can be seen, the wind autocorrelations are linear combinations of the functions f and g when $N_{\psi\psi} = N_{\chi\chi}$. If the errors are assumed to be nondivergent ($\nu^2 = 0$), f and g are the autocorrelation functions of the longitudinal and transverse components, respectively. These functions are also shown in Figure 27.

Finally, height-wind correlations are obtained from the relation

$$\rho_{z\tilde{v}} = L\sqrt{1-\nu^2}\frac{d}{dr}\rho_{z\psi} \ ,$$

where it was assumed that $\rho_{z\chi} = 0$, from which it follows that $\rho_{z\tilde{u}} = 0$. In the baseline version of the TAP error statistics, height and streamfunction are geostrophically coupled with an adjustable parameter μ ($|\mu| \leq 1$): $\rho_{z\psi} = \mu\rho_{zz}$.

An example of the resulting horizontal correlation functions in terms of the local (east/north) coordinate system is given in Figure 28. For those plots, parameters appropriate for a tropical location were chosen ($\mu = 0.5$ and $\nu^2 = 0.1$) – the map background is only provided to give a sense of the spatial scale.

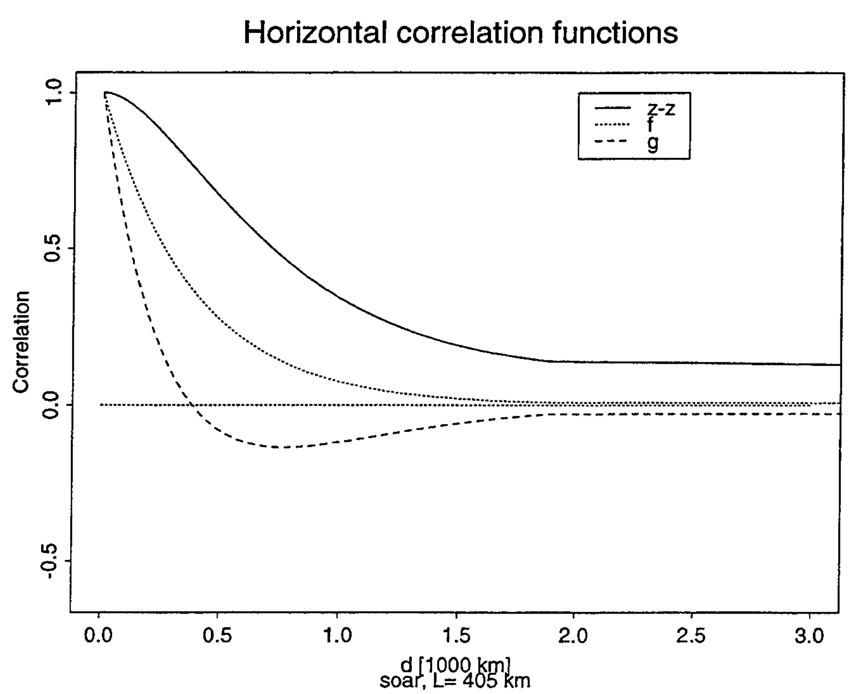


Figure 27: Horizontal height error correlation functions for NOGAPS (see text). The horizontal axis is separation distance in units of 1000 km

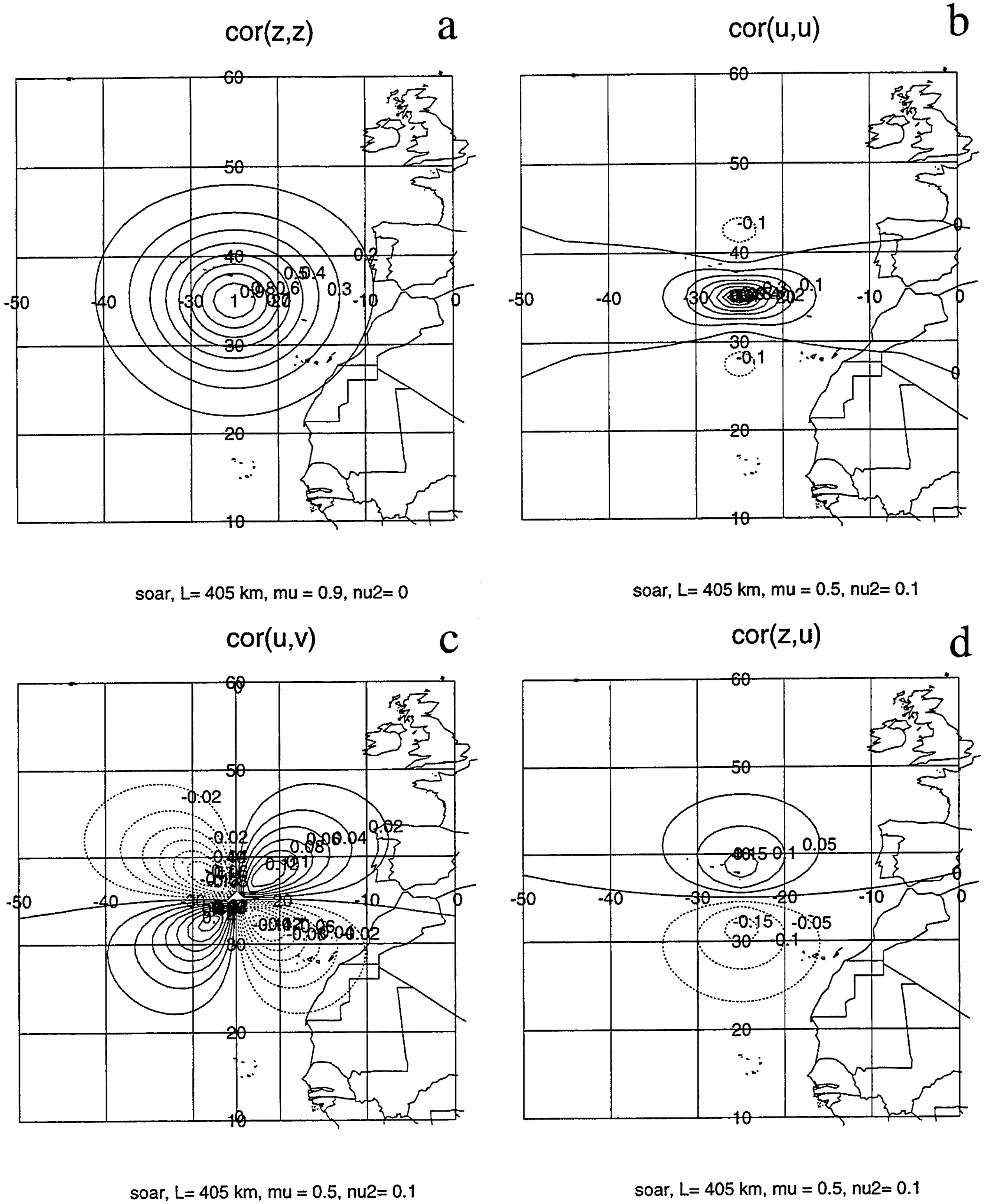


Figure 28: Horizontal maps of NOGAPS forecast error correlations with a single observation at the grid center: (a) height-height; (b) zonal wind - zonal wind; (c) zonal wind - meridional wind; (d) height - zonal wind.

For the climatology background, the horizontal error correlation function is given by

$$\rho_{zz}(r) = \frac{\alpha\cos(\omega r) + 1 - \alpha}{\sqrt{(1 + \lambda^2 r^2)}} ,$$

with $\alpha = 0.738$, $\omega = 1.38 \cdot 10^{-3} km^{-1}$, and $\lambda = 0.848 \cdot 10^{-3} km^{-1}$. The function is shown in Figure 29, along with its rescaled first and second derivatives. The resulting horizontal maps of height-height, wind-wind, and height-wind correlations are shown in Figure 30.

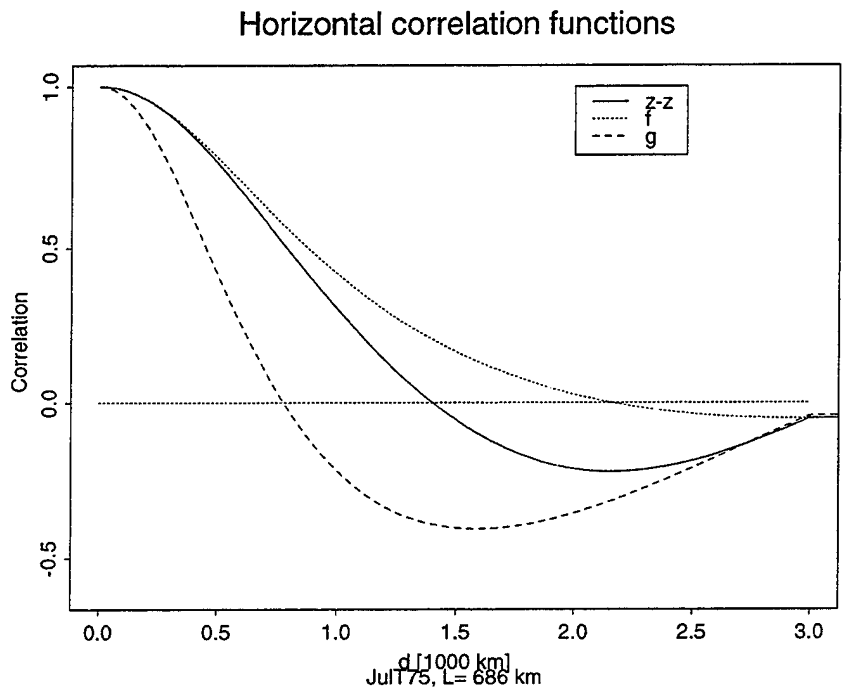


Figure 29: Horizontal height error correlation functions for climatology.

Finally, the relative humidity error correlations, for NOGAPS and climatology backgrounds, are modeled by a negative squared exponential (NSE) function:

$$\rho(r) = e^{-0.5(r/L)^2} ,$$

with $L = 300km$. This function is shown in Figure 31.

4.3.3 Vertical correlations

For the baseline version of the TAP error statistics, the vertical correlation functions of the velocity potential and streamfunction autocorrelation, and the cross-correlation of height with the transverse wind component, are set equal to the height-height autocorrelation function. The functional form given in Goerss and Phoebus (1993 [GoeP93]) is

$$N_{zz}(z_1, z_2) = exp(-(\frac{|z_1 - z_2|}{dz})^b) ,$$

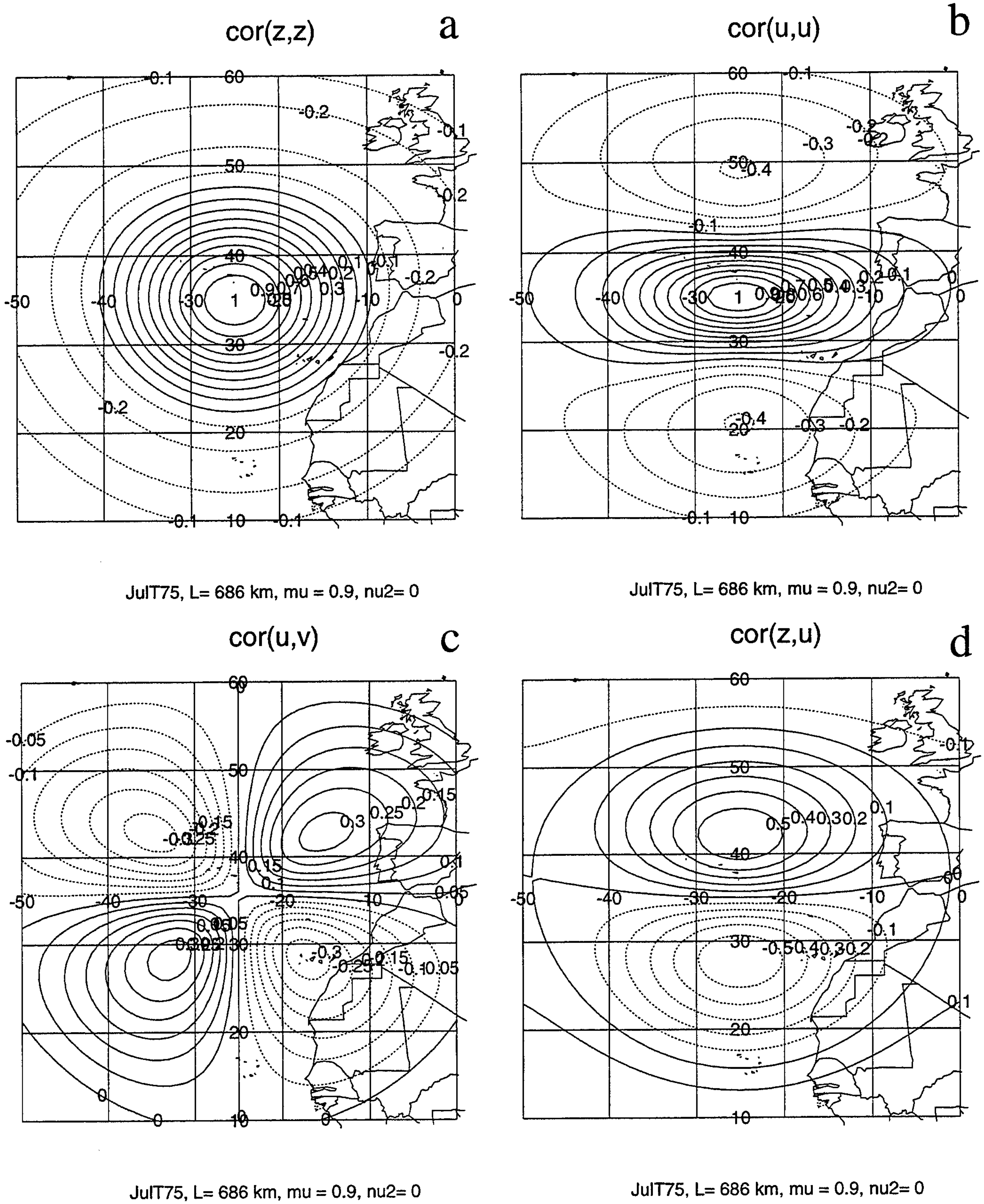


Figure 30: Horizontal maps of climatology error correlations with a single observation at the grid center: (a) height-height; (b) zonal wind - zonal wind; (c) zonal wind - meridional wind; (d) height - zonal wind.

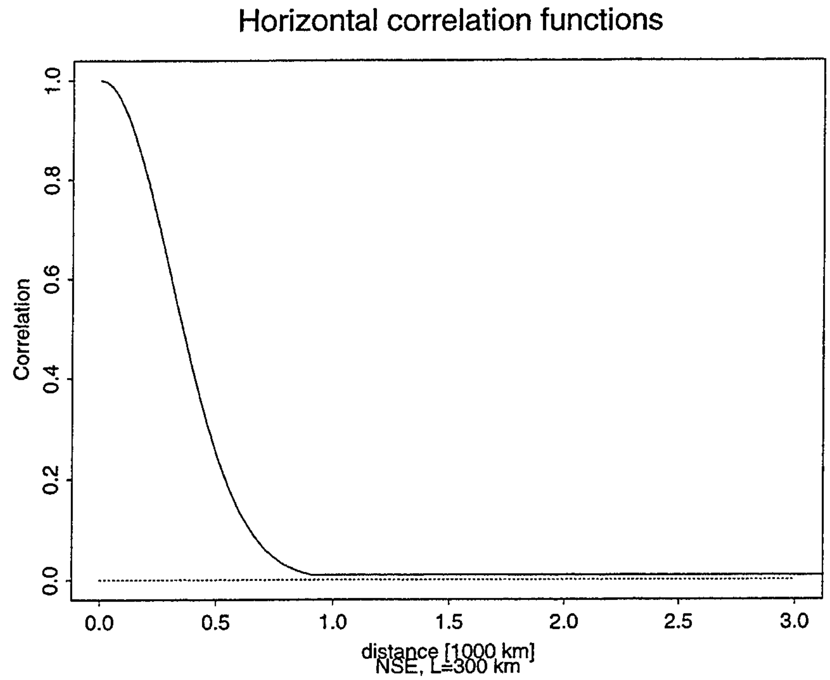


Figure 31: Horizontal relative humidity error correlation functions for NOGAPS and climatology.

where z_1 and z_2 are the height of the two observations, b and dz are adjustable constants ($b = 1.8$, $dz = 3600\ m$). As used in TAP, the natural logarithm of pressure is used instead of height, and the length scale dz is replaced by the equivalent log-pressure scale obtained from the hydrostatic relationship

$$|d \ln p| = \frac{g}{RT} |dz| ,$$

where g is the acceleration of gravity, R the gas constant for air, and T a reference temperature ($T = 250\ K$). The functional form for N_{zz} is evaluated for all possible combinations of the standard pressure levels and stored in tabular form. Figure 32 shows the resulting correlations for an observation at 500 hPa, both from the continuous functional form, and as evaluated by interpolation from the table.

The tabular data of N_{zz} are further used to derive the vertical correlation table for temperature. For any two variables that are related linearly, as in

$$\boldsymbol{T} = \mathrm{M}\boldsymbol{z} ,$$

where $\boldsymbol{T}$ and $\boldsymbol{z}$ are column vectors containing temperature and height values at the N pressure levels, and M is an N by N matrix, the covariance matrix of $\boldsymbol{T}$ ($S_T = \langle \boldsymbol{T}\boldsymbol{T}^* \rangle$) can be computed as follows:

$$S_T = \mathrm{M}\langle \boldsymbol{z}\boldsymbol{z}^* \rangle \mathrm{M}^* = \mathrm{M} S_z \mathrm{M}^* ,$$

where M^* indicates the matrix transpose. The matrix M is obtained from the application of the hydrostatic relationship. In these computations, the tabulated values of the height standard deviations are used to convert N_{zz} to S_z. The resulting values of N_{TT} for a temperature observation at 500 hPa are also shown in Figure 32.

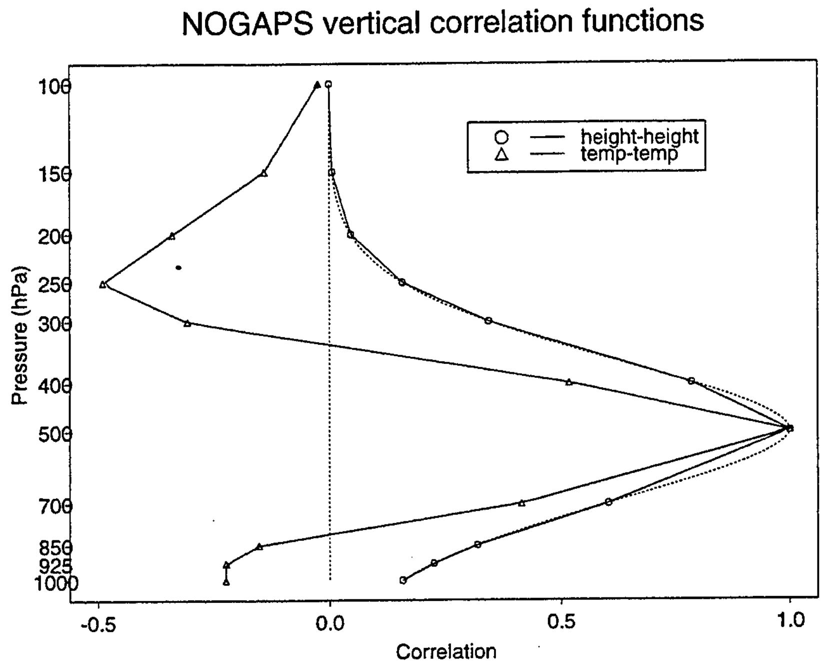


Figure 32: Vertical error correlation functions for NOGAPS for an observation at 500 hPa. Values from table shown as symbols. At intermediate pressures, plotted values are interpolated from the table (solid lines) or computed from continuous functional form (dashed line)

For climatology backgrounds, no published sources for vertical correlation functions have been found. Computation from available data sources has not been performed to this date because of problems of assembling and processing the required data bases. A standard radiosonde dataset we had planned on using for this purpose (the so-called TIGR dataset used in connection with satellite retrievals) turned out to be unsuitable because it did not contain date/location information for the individual profiles. Because it was then impossible to separate contributions from geographical versus time variations to the covariances, the vertical correlations of the deviations from the overall mean profile are unrealistically large when applied to a gridded climatological first-guess field. Acquisition and quality control of other archived radiosonde data was postponed for later phases of the project. Instead, we generated vertical correlations from the same functional form as for the NOGAPS data, but adjusted the function parameters for somewhat broader correlation functions ($dz = 6500\,m$). The resulting height-height and temperature-temperature correlations for a 500 hPa observation are shown in Figure 33.

The only vertical correlation functions of observation errors currently used in TAP are those of radiosonde heights. They are taken directly from Lönnberg *et al.* (1992 [LonSU92]). These values are shown in Table 2.

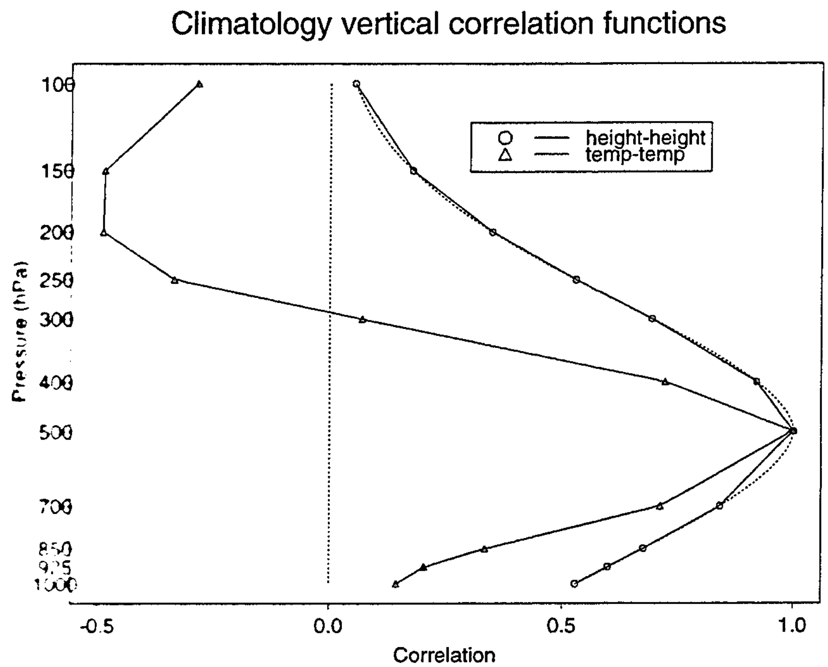


Figure 33: Vertical error correlation functions for climatology for an observation at 500 hPa. Values from table shown as symbols. At intermediate pressures, plotted values are interpolated from the table.

Table 2: Vertical correlation ($\times$1000) of radiosonde height observation errors used in TAP.

	1000	850	700	500	400	300	250	200	150	100	70	50	30	20	10
1000	1000	716	276	29	5	1	0	0	0	0	0	0	0	0	0
850	716	1000	733	183	55	9	3	1	0	0	0	0	0	0	0
700	276	733	1000	573	268	77	31	10	2	0	0	0	0	0	0
500	29	183	573	1000	851	480	288	138	49	11	3	1	0	0	0
400	5	55	268	851	1000	814	601	364	167	51	19	7	1	0	0
300	1	9	77	480	814	1000	935	738	458	200	93	44	11	2	0
250	0	3	31	288	601	935	1000	919	678	361	194	104	31	7	0
200	0	1	10	138	364	738	919	1000	895	597	375	229	84	24	2
150	0	0	2	49	167	458	678	895	1000	861	649	460	214	78	9
100	0	0	0	11	51	200	361	597	861	1000	929	782	480	230	42
70	0	0	0	3	19	93	194	375	649	929	1000	951	710	412	103
50	0	0	0	1	7	44	104	229	460	782	951	1000	878	598	192
30	0	0	0	0	1	11	31	84	214	480	710	878	1000	881	426
20	0	0	0	0	0	2	7	24	78	230	412	598	881	1000	725
10	0	0	0	0	0	0	0	2	9	42	103	192	426	725	1000

5 Data Collection for Real-Data Tests

5.1 Data Sources and Collection Procedures

Real data for system tests have been collected from a variety of sources. The selection of the data sources was driven by the competing requirements of ease of access and ingest of the data on the one hand, and adequate geographical coverage and representation of different background and data types on the other. Conventional (surface, ship, buoy, and upper air) data were collected and decoded using the Family of Services database and associated software on the AIMS (Air Force Interactive Meteorological System) system located at the Phillips Laboratory and operated by AER personnel. Direct-readout satellite data were collected from the NOAA and DMSP polar orbiters, using the AIMS ground station equipment. The direct-readout satellite data cover only the area visible from the satellite while transmitting to the ground station, which is centered over the East Coast of the United States. Because of this restriction, conventional data were only collected between latitudes 15° N and 60° N, and longitudes 45° W and 105° W. A variety of gridded model forecast and analysis fields were obtained from the anonymous NCEP ftp-server. The gridded fields obtained from NCEP were all in GRIB format, and software was assembled and/or written to decode these data and ingest them into the TAP analysis system. The satellite direct-readout data was processed using the TOVS export package for obtaining temperature and moisture retrievals from the NOAA polar orbiter data software, which was installed on the AIMS ground station. In the tests reported here, the DMSP data were not used. In addition, aircraft reports were obtained in an ASCII format by PL personnel from data archives at ETAC, and software was written to ingest them into the TAP analysis system.

5.2 Cases Collected

In all, data for four separate case study days have been archived: 6-7 March 1995, Hurricane Erin (2-3 August 1995), Hurricane Opal (4-5 October 1995), and 7-8 March 1996. A brief description of all four cases is given here.

5.2.1 6-7 March 1995

Data have been collected for approximately a 24-hour period centered around 12 UTC 6 March 1995. The synoptic situation on that day was characterized by an upper-level shortwave passing through the Northeastern United States and Canada, embedded in a southwesterly upper-level flow. At the surface, this was accompanied by a weak low (central pressure around 1016 hPa) and rain over the Northeastern United States, snow over parts of Quebec and the Canadian Maritimes. This case represents a typical moderate to weak winter/early spring storm over the Northeastern United States. Sample TAP plots from this case are shown in Appendix B.

5.2.2 Erin: 2-3 August 1995

A summertime case was selected to coincide with the landfall of hurricane Erin in Florida. Data were collected for the period 09 UTC 02 August 1995 – 03 UTC 03 August 1995.

While most of the East Coast was dominated by high pressure and weak flow at all levels, hurricane Erin made landfall in central Florida (near Vero Beach) at 0545 UTC 2 August. While over the Caribbean waters, this hurricane had maximum sustained winds of 85 mph, and a central pressure as low as 980 hPa. After landfall, winds dropped to below 70 mph and it was downgraded to Tropical Storm status. At 12 UTC, it was roughly centered over the Florida peninsula, and by 20 UTC its center had moved over the Gulf waters. It subsequently reintensified to hurricane strength and made landfall in the Florida panhandle (in Pensacola Beach) at 1530 UTC 3 August. A secondary feature of interest on this day are the remnants of tropical storm Dean, which were located over parts of Texas, Oklahoma, and Kansas, and which had caused widespread convective rain and flooding.

5.2.3 Opal: 4-5 October 1995

This case coincided with the landfall of Hurricane Opal in the Florida panhandle (right near Hurlburt Field). The time period covered by our archive is 1800 UTC 4 October – 1800 UTC 5 October 1995. Opal made landfall at approximately 2300 UTC 4 October. Opal was a strong hurricane, causing over $3 billion of damage and over 20 deaths. It thus represents an extreme weather event. Unfortunately, because of the extreme weather conditions, key radiosonde reports are missing for this case at some or all of the levels. Over the Northeast United States, there are weak shortwave features embedded in a generally southwesterly flow. Sample TAP plots from this case are shown in Appendix B.

5.2.4 7-8 March 1996

Data has been collected for 00 UTC 7 March – 12 UTC 8 March 1996 for this case. Over this period, a surface low moved from the Georgia/North Carolina border northeastward to the east of Massachusetts. At the beginning of the time period, the Southeast United States was experiencing strong convective activity with this system, while at later times the Northeastern seaboard was affected with extensive areas of precipitation, falling as rain to the south and snow to the north of a snow/rain boundary located roughly in central Pennsylvania. This case was used to test the version of TAP delivered at the end of the contract. Results are shown in Section 7.

6 Early Prototype Tests at the Combat Weather Facility

6.1 User Interface

While development of a complete user interface and display capability is beyond the scope of the contract (these components are assumed to be available on the host system), these capabilities had to be provided in a united form for testing the early prototype system. For system development and testing, we developed a programming/execution environment which makes use of various tools available to us in the Unix environment (Splus, Gnu Make, Gnu Emacs). Because the test personnel had little or no familiarity with the Unix operating system, the Gnu Emacs editor, or the Splus programming language, the user interface for the real data tests was designed to enable users to run the TAP system from end-to-end with a few simple commands, allowing choices of only a small number of preselected options. Two separate user interfaces were developed: a cshell interface, consisting of a set of cshell scripts to be run from the Unix command prompt, and a HyperText Markup Language (HTML) interface which uses a Web browser (such as Netscape) for user interaction and the display of the results. Both user interfaces are fully described in the User's Manual produced for the real data tests (see Appendix B).

6.2 TAP Setup

As was mentioned above, only a small subset of all possible configuration options was made available for the real data tests. Two of the four real cases were provided for testing (the March 1995 and Opal case), and two areas of the country were selected as possible analysis areas. Within each area, one of three possible grid configurations could be selected: The outermost, large grid domain covers a region of approximately 1500 km on a side. Centered inside the outer region is the small domain, covering a region of approximately 750 km on a side. Both grids consist of 11 by 11 gridpoints. Finally, the column domain represents the grid column at the center grid point of the small analysis domain. Over the Northeast region, a polar stereographic projection basemap is used (with a reference longitude at 80° W), and the analysis domains are centered over southeast Pennsylvania (see Appendix B). Over the Southeast region, a Mercator projection basemap is used, and analysis domains are centered over Alabama (see Appendix B). The analysis grids were chosen to contain both land areas with dense data coverage and data sparse areas over water. The relatively small grid (11 by 11 gridpoints) was chosen to enable the early prototype code to execute in a timely manner on the workstation used for testing.

Users could choose to perform a surface temperature analysis, or an upper air analysis at a set of preselected pressure levels, for temperature, height, winds, height and winds, or relative humidity. Radiosonde data were used for the upper air analyses, and surface reports for the surface analyses. Either climatology or a 12-hour ETA model forecast could be used for the background field. The resolution of the climatology background (5° longitude by 2.5° latitude) is significantly coarser than even that of the outer analysis grid used in the tests (150 km), whereas the ETA model forecast are available at only slightly coarser resolution (190.5 km at 60° N).

6.3 TAP Installation

We installed the TAP software and data, along with the needed supporting software, on a Sun workstation at Phillips Laboratory (PL) during April and May. During this exercise, we discovered and corrected numerous minor problems with the installation procedure. However, because of compiler incompatibilities between the PL machine and the Splus software (and the Unix installations at AER and CWF), the installation could not be completed at PL. The TAP installation was then repeated on June 4 and 5 at CWF in Hurlburt Field, Florida. With the support of AER computer system staff, and the CWF personnel on-site, both the TAP software and the supporting software (including Splus, Netscape, and Sun compilers) were successfully installed on the Sun workstation at CWF. (A missing memory module on the workstation, however, resulted in somewhat slower execution times.) A written User's Manual (Appendix B) and oral training were provided by AER to enable the test personnel to run and evaluate the early TAP prototype.

6.4 Results of the Prototype Tests

The real-data testing was performed by CWF during June and July, and results (in the form of completed test questionnaires, and summary comments) were returned to PL. These results are summarized below.

A total of eight testers generated between 1 and 3 TAP analyses each. None of the testers had any prior experience with TAP, but most had prior experience with weather analysis and forecasting, and rated their experience with weather displays, and understanding of forecast products, highly (with a score of 3 or higher on a scale from 1 to 5).

The setup and running of TAP presented few serious difficulties for the testers. Five of the eight testers read at least part of the User's Guide, and did not find the instructions difficult. All but one of the testers thought later exercises were easy after the first successful completion: successful completion was obvious, and instructions for the display of the results were clear. Six of the eight testers followed the diagnostic messages, and they all found them understandable. Only one tester experienced an abnormal stop (none of the displays would work). Slightly over one-half of the testers reported some problems (one in setup, one in execution, and three in display); however, all but one tester were able to obtain the expected results from the TAP execution on the first try.

Execution time averaged 5 minutes for the setup, 9 minutes for TAP execution, and 1-2 minutes for display (with the exception of one reported 30 minute time for display). These execution times are generally in line with our own timing tests, and the estimates provided as part of the user interface and User's guide.

The opinion on the realism of the TAP analyses and their usefulness for forecasting was split: 3 testers rated the realism poorly (1 on a scale of 1-5), 2 rated their usefulness for forecasting poorly (1 or 2), but all others gave fair to good marks in both categories (3 or 4). Opinions were similarly divided on whether the data available for the analyses limited their realism and usefulness. A wide range of opinion (from 1 to 5) existed on whether TAP was inflexible (average response = 3.4) or whether it was hard to use (average=3.2).

Of the written comments provided in the questionnaires, and the overall comments by CWF, the most serious criticisms concerned the level of development of TAP: the restriction

to predefined analysis regions, which were of small extent and coarse resolution, the restricted flexibility with respect to data sources, and the slow execution speed. The limited possibilities for intercomparison of TAP analyses with observational data and reference analyses was also criticized. Other comments and suggestions for improvements centered on areas outside the scope of the main TAP development effort: desired options for the display of derived quantities (such as vertical velocity, vorticity), different display methods (e.g., wind barbs or streamlines instead of wind vectors), and different units (knots instead of m/s, °F or °C instead of K).

6.5 Discussion

The setup, installation, user interface and documentation for the real data tests of TAP all fulfilled their stated purpose: to enable an end-to-end test of the early prototype system by personnel with no prior experience with TAP. The fact that even those testers that did not read the User's guide successfully completed their exercises on the first try proves the user-friendliness of the user interface.

Clearly, however, the test results also pointed out areas for improvement for TAP. Perhaps most important is the need to increase the execution speed (and improve the memory management), to enable TAP to complete analyses on larger analysis grids in a timely manner. This is in fact one of the main system development tasks in the final year of the project. The other main area of improvement is the need to support additional data types, aside from those used in the prototype tests. The real-data tests thus served one of their main objectives: to provide user feedback during the development cycle, and help to focus the development effort on the most important aspects of the analysis system.

The negative ratings with respect to analysis realism and usefulness appear to be due to restrictions of the prototype system that are a direct result of the development stage or testing setup, and not representative of the final operational product. In particular, the restrictions in grid size, availability of observational data, and other TAP options all were dictated by the logistical difficulties of testing an early prototype in a quasi-operational environment. In addition, since the only tool available for evaluation of the analyses was their display, their usefulness for initialization of forecast models or input to other weather support algorithms could not be fully assessed.

7 Prototype Tests for Mesoscale Model Initialization

7.1 Experimental Setup

In preparation for the installation and evaluation of TAP at the AFWA, where TAP is to be used for initialization of the MM5 mesoscale forecast model, a number of tests were performed to demonstrate the suitability of TAP for this purpose.

For this application, TAP is inserted in the MM5 preprocessing procedure. In the usual MM5 preprocessing procedure, the model grid domain and certain fixed fields are prepared by running program `TERRAIN`; in the next step, a large-scale gridded field (such as a global analysis or forecast) available on pressure surfaces is interpolated horizontally to the model gridpoints by running program `DATAGRID`; a successive corrections analysis program (`RAWINS`) is then used to modify the `DATAGRID` output, based on radiosonde observations; finally, the pressure level fields are interpolated vertically to the MM5 model levels by program `INTERP`. Some additional variable transformations and initialization steps are performed by `INTERP`, as well.

In the tests reported here, TAP is inserted in the MM5 preprocessing in place of the `RAWINS` program, *i.e.* TAP performs a pressure-level analysis (of geopotential height, temperature, relative humidity, and winds) at the horizontal MM5 grid points. The background (first guess) information was provided either by the ETA model 12-hour forecasts archived at AER as part of the real data collection described earlier, or by NCEP global analysis fields archived at NCAR. In the case of the global analysis background, the `DATAGRID` program was used to interpolate the global analysis to the MM5 grid chosen for these tests (see below). The file created by `DATAGRID` was then ingested into the TAP preprocessor to provide the information needed to determine the TAP analysis grid, levels, and variables, and the values of the background field. In the case of the ETA model forecast background fields, the analysis grid, levels, and variables were set up (by the appropriate TAP functions) to match the DATAGRID output, and the existing TAP ingest routines for GRIB format data files were used to provide the values of the background field. In either case, the output of TAP is then postprocessed to the same format as the DATAGRID output, so it can be used as input to the INTERP program.

We performed 24-hour forecasts for the 7-8 March 1996 case, using TAP analyses for the 12 UTC 7 March time period as the initial state. The analysis grid, which is also the MM5 forecast model grid, is a 75x65 grid on a polar stereographic projection with a grid spacing of 50 km. The analysis levels were at 1000, 850, 700, 500, 400, 300, 250, 200, 150, and 100 hPa. The MM5 was run with 30 layers in the vertical, with vertical levels at $\sigma =$ 1.000, 0.992, 0.980, 0.966, 0.950, 0.934, 0.918, 0.902, 0.886, 0.866, 0.842, 0.814, 0.780, 0.740, 0.694, 0.648, 0.600, 0.556, 0.510, 0.464, 0.418, 0.372, 0.326, 0.282, 0.240, 0.198, 0.156, 0.114, 0.074, 0.036, 0.000.

We used version 2 of the MM5 mesoscale model (Grell *et al.*, 1994 [GreDS94]) in nonhydrostatic mode with explicit prediction of water vapor, cloud and rain water mixing ratios. The ground temperature is also predicted from a computed surface energy budget. Other physical processes are parameterized in these simulations, including convective fluxes, long- and short-wave radiation, turbulent boundary layer mixing and warm- and cold-cloud precipitation processes.

MM5 forecasts out to 24 hours were produced for four separate cases (see Table 3), which differed in the type of background field, and whether or not observations were provided to TAP to modify the background field. In all cases, lateral boundary conditions were provided from global NCEP analyses (processed through DATAGRID) throughout the length of the forecast. In the case with observations, radiosonde reports, TOVS retrievals, and aircraft reports were used in the TAP upper air analysis. The ETAnotap and ETAtap runs serve to demonstrate the suitability of the TAP output for model initialization, for the case when TAP is run with a short-term forecast as a background field. These runs use the configuration tested in the previous real-data tests, *i.e.* using the TAP functions for setup and ingest of the GRIB format operational short-term forecast, and TAP functions for ingest of the AIMS format observations. Runs GLOtap and GLOnotap provide a test of the preprocessing and ingest routines written to make use of the DATAGRID output for analysis grid setup and background field initialization. Comparison of the runs with and without observations serve to demonstrate the impact of using TAP to refine the model initial state.

Table 3: MM5 test cases

Name	Background	Observations
ETAnotap	ETA 12-hour forecast	No
ETAtap	ETA 12-hour forecast	Yes
GLOnotap	NCEP global analysis	No
GLOtap	NCEP global analysis	Yes

7.2 Analysis Results

The initial state 500 hPa height and winds of the ETAnotap run (Figure 34), which is the ETA model 12-hour forecast, shows an upper-level trough in the lee of the Rocky Mountains, with generally southwesterly flow over the East coast. Comparison with the corresponding ETAtap analysis (Figure 35) shows that the TAP analysis lowers heights over the entire analysis domain, most notably from Illinois northeastward toward New England, which acts to deepen and sharpen the upper level trough. The corresponding figures for the 850 hPa level (Figures 36 and 37) show a deepening of the trough, which is most pronounced over the Northeast states at that level.

It is interesting to compare these results with the GLOnotap and GLOtap initial state fields, since the GLOnotap is already an analysis field, albeit from a large-scale analysis. At 850 hPa (Figure 38), the GLOnotap field bears a closer resemblance to the ETAtap analysis than the ETA forecast, as is to be expected. Consistent with this, the analysis increments are smaller for GLOtap than ETAtap, and the GLOtap analysis (Figure 39) is quite similar to the corresponding background field.

The situation is somewhat different for the relative humidity fields. While both the ETAnotap (Figure 40) and GLOnotap (Figure 41) fields at 850 hPa exhibit the main features of this case (a large moist area along and ahead of the trough, and generally drier air behind it), there are substantial differences in some of the details. The GLOnotap-ETAnotap

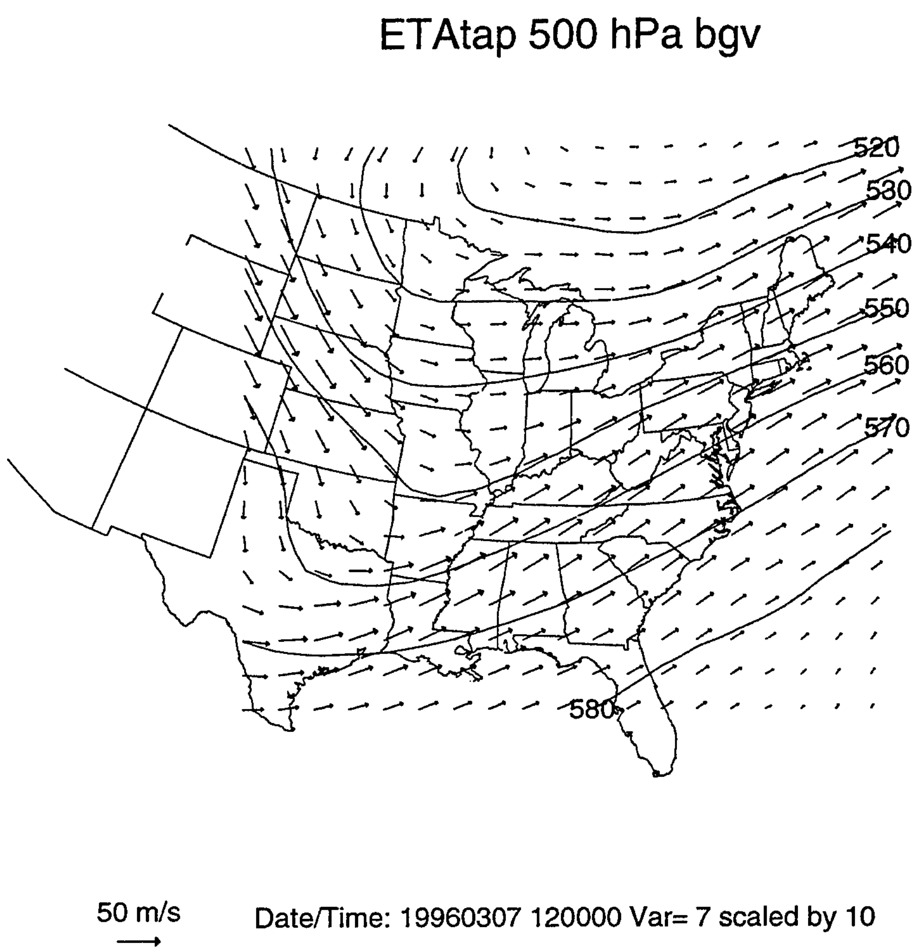


Figure 34: ETAnotap initial state: 500 hPa height (dam) and winds from the 12-hour ETA model forecast.

ETAtap 500 hPa anv

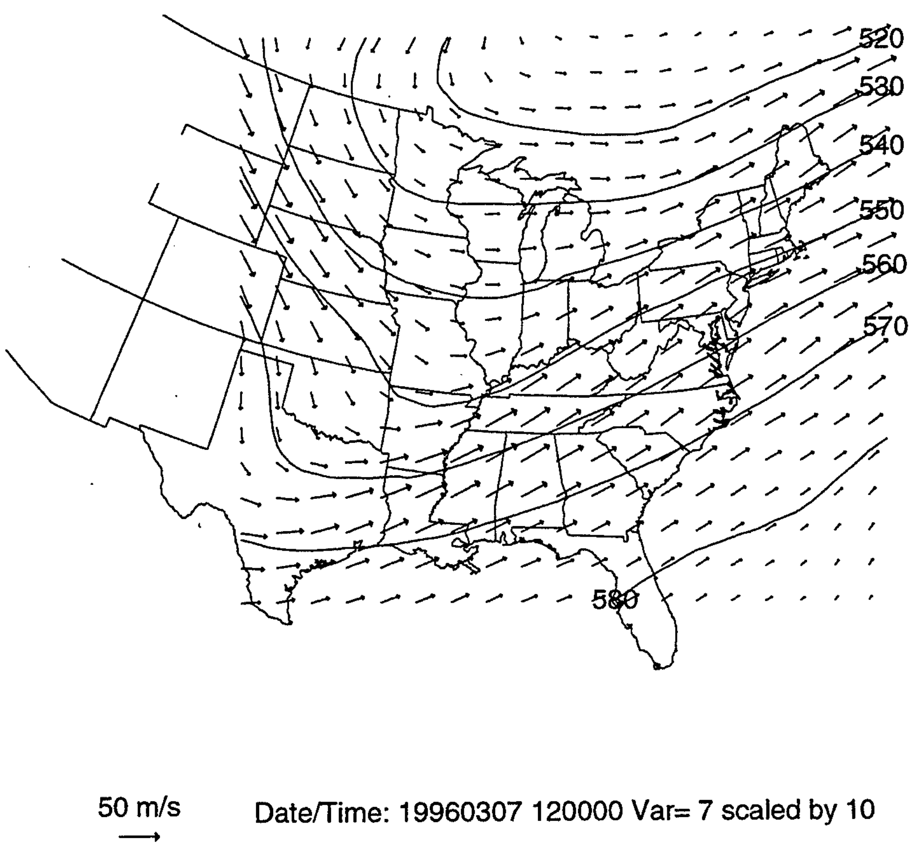


Figure 35: ETAtap initial state: 500 hPa height (dam) and winds from the TAP analysis.

ETAtap 850 hPa bgv

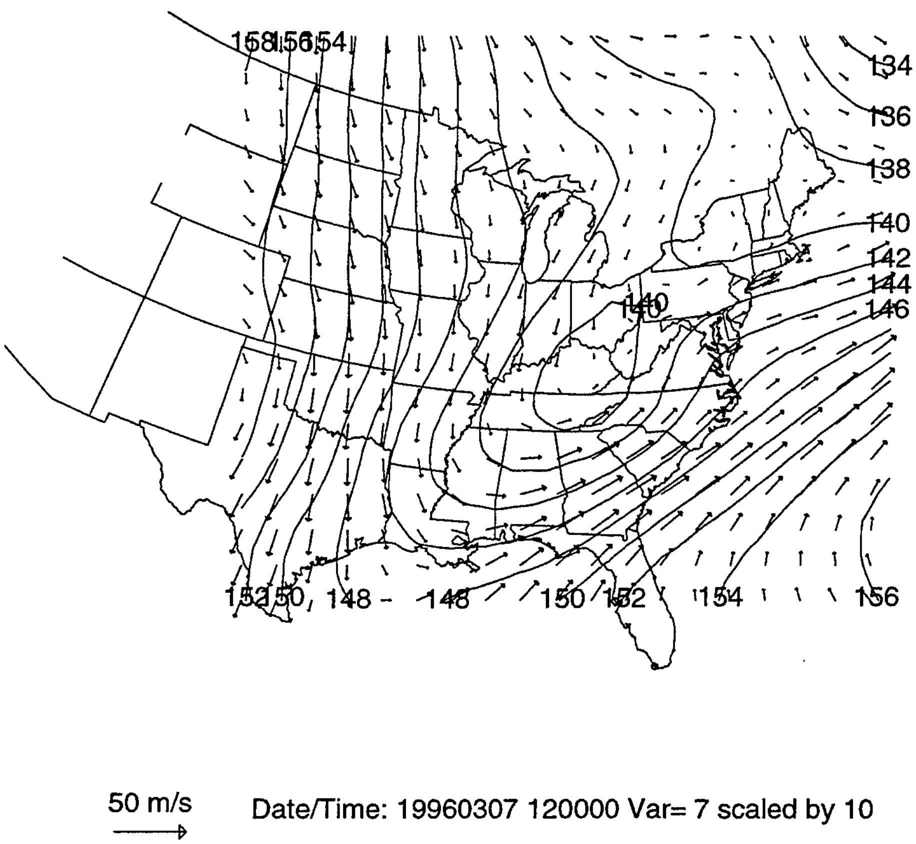


Figure 36: ETAnotap initial state: 850 hPa height (dam) and winds from the 12-hour ETA model forecast.

ETAtap 850 hPa anv

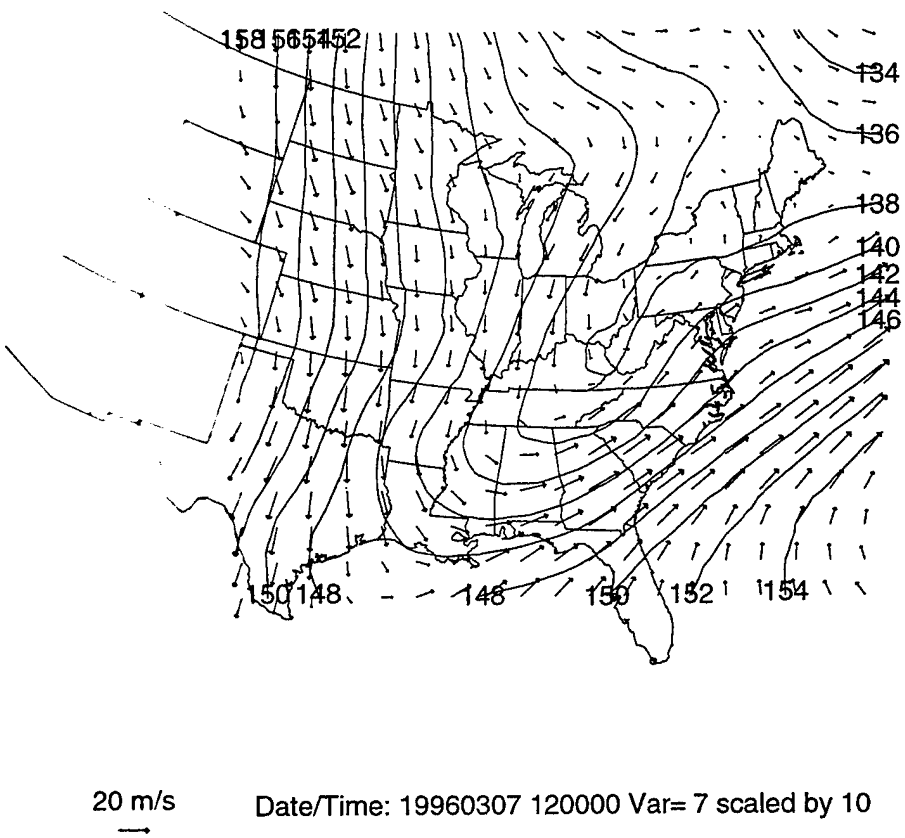


Figure 37: ETAtap initial state: 850 hPa height (dam) and winds from the TAP analysis.

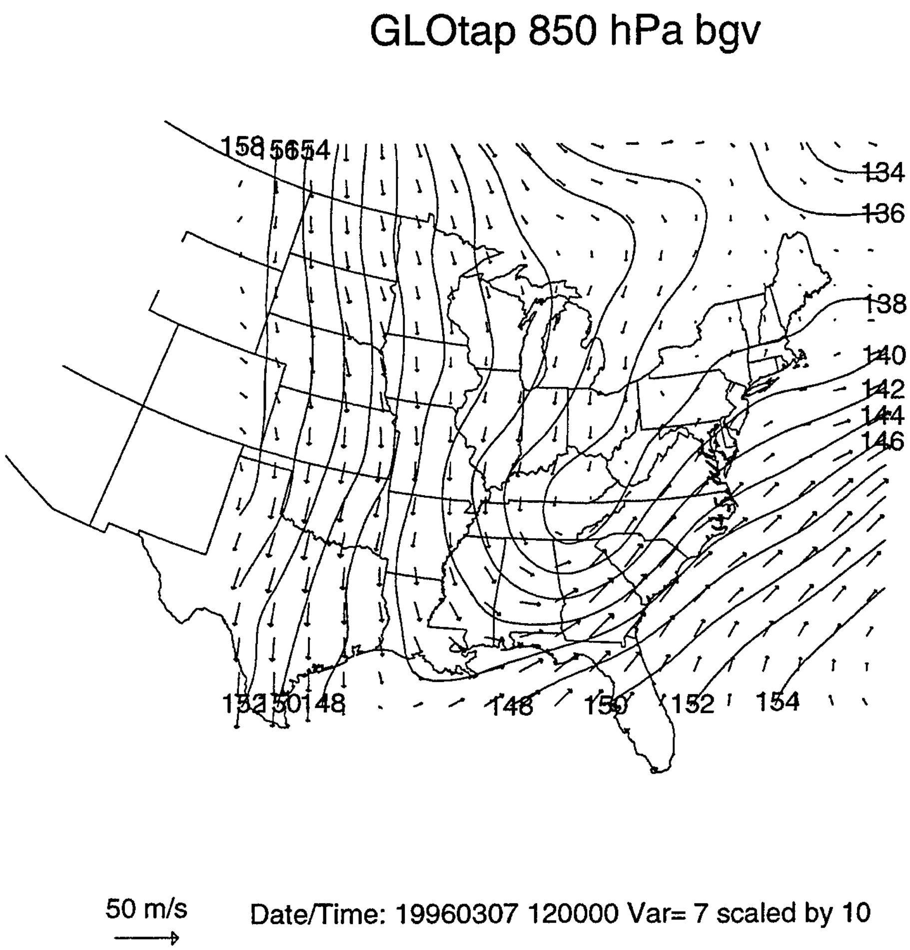


Figure 38: GLOnotap initial state: 850 hPa height (dam) and winds from the global analysis.

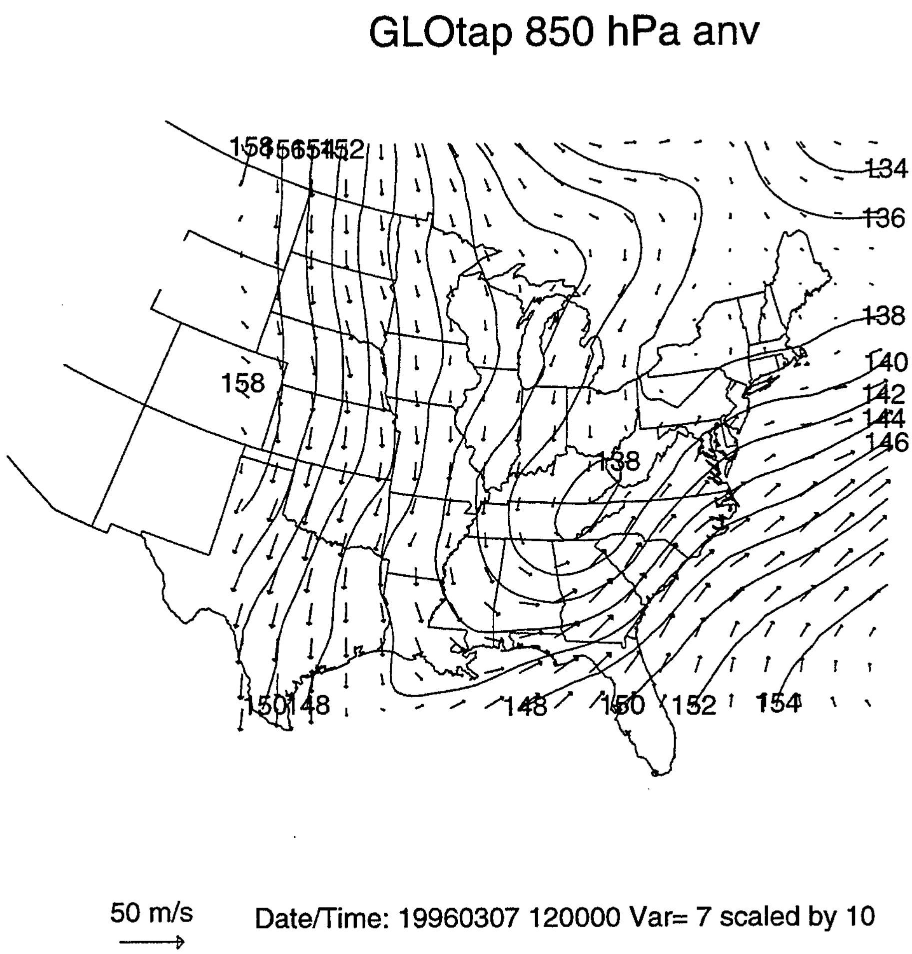


Figure 39: GLOtap initial state: 850 hPa height (dam) and winds from the TAP analysis.

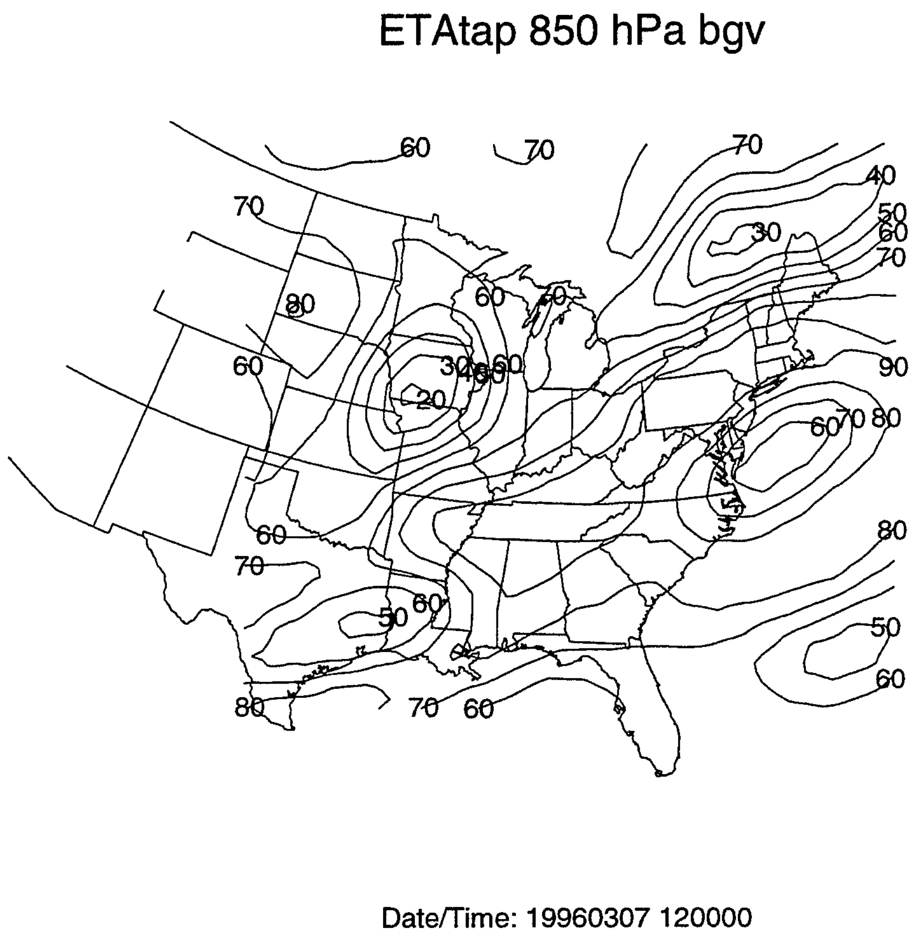
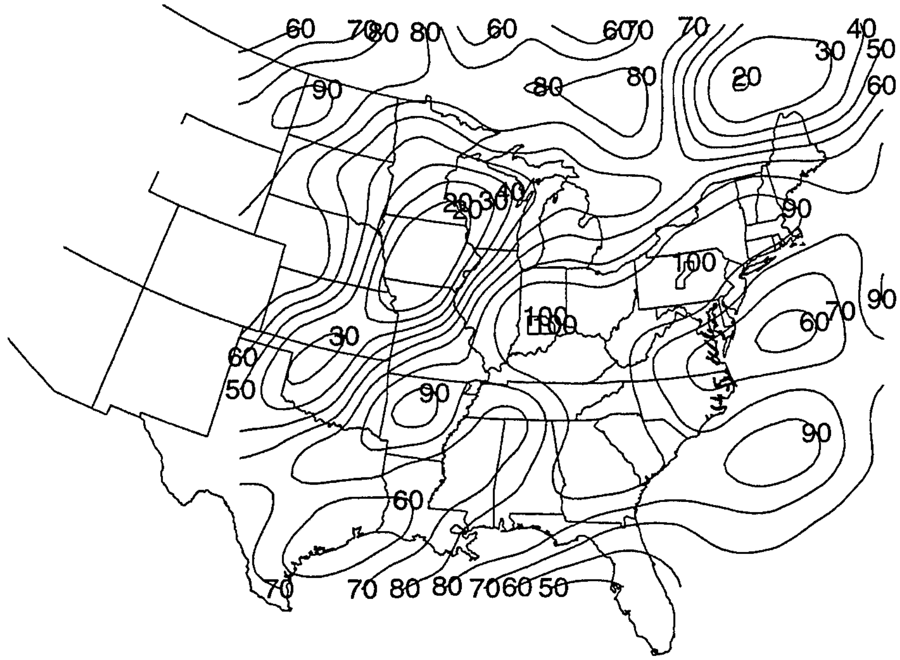


Figure 40: ETAnotap initial state: 850 hPa relative humidity from the 12-hour ETA model forecast.

GLOtap 850 hPa bgv



Date/Time: 19960307 120000

Figure 41: GLOnotap initial state: 850 hPa relative humidity from the global analysis.

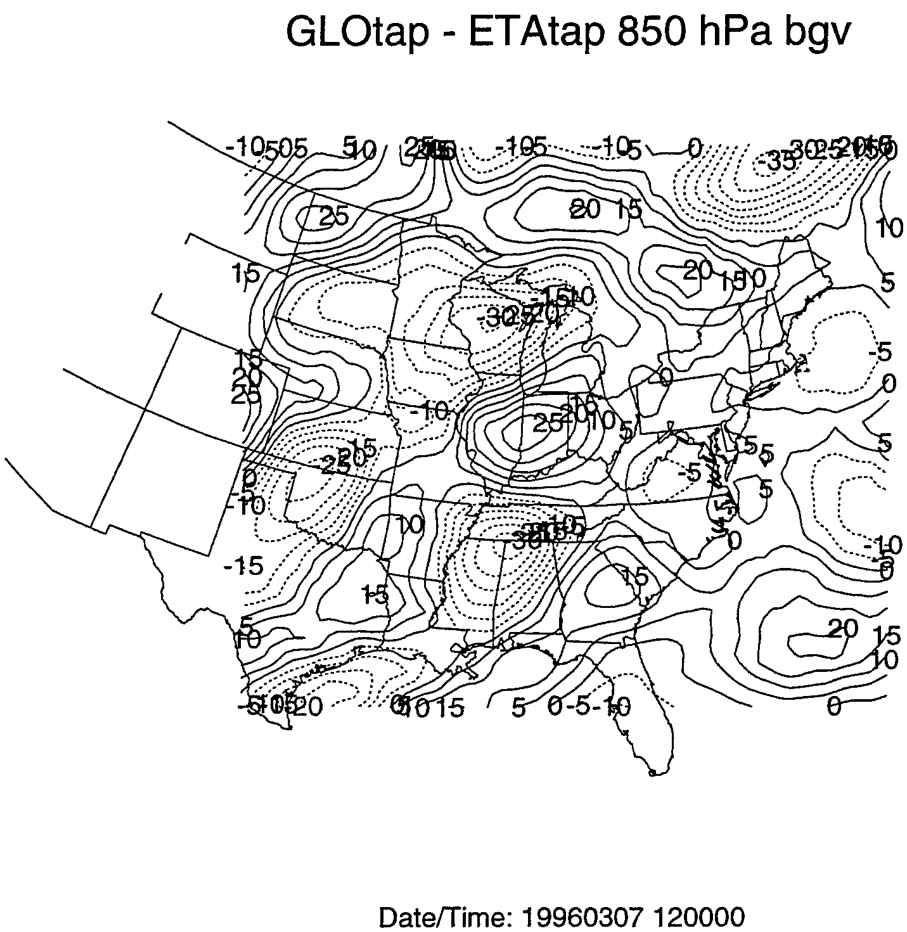
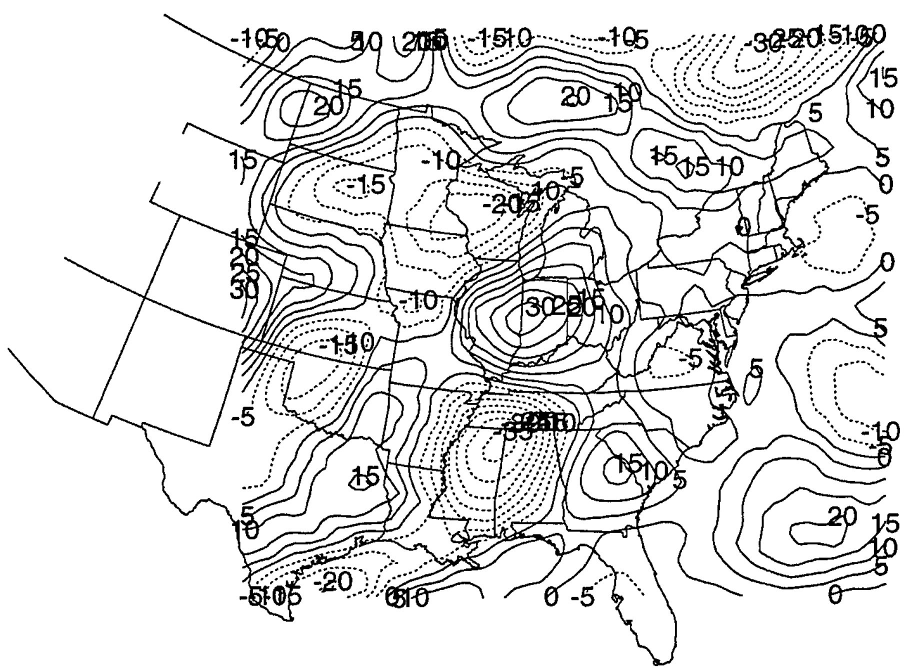


Figure 42: Difference between GLOnotap and ETAnotap initial state: 850 hPa relative humidity.

GLOtap - ETAtap 850 hPa anv



Date/Time: 19960307 120000

Figure 43: Difference between GLOtap and ETAtap initial state: 850 hPa relative humidity.

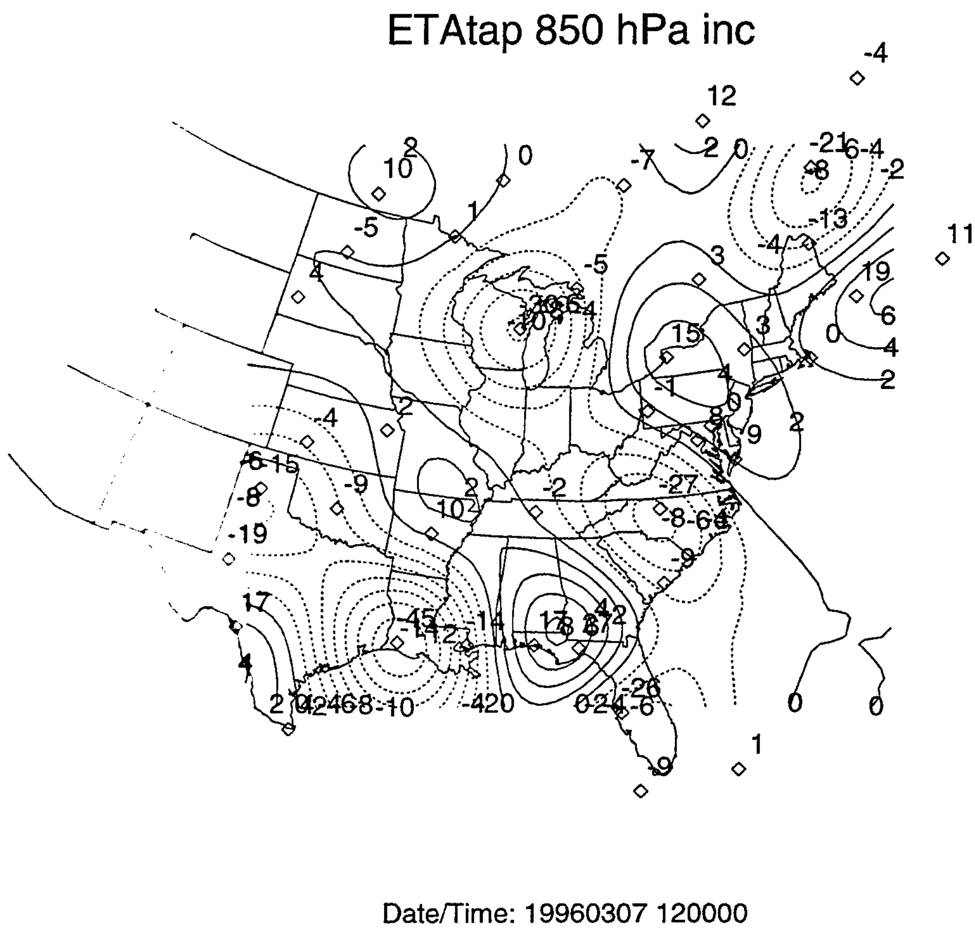


Figure 44: ETAtap analysis and observation increments of 850 hPa relative humidity. Analysis–background values shown as contours, radiosonde observations–background values plotted at the observation locations.

difference field (Figure 42) contains small-scale features with large amplitudes of up to 35%. Interestingly, the corresponding difference field for GLOtap-ETAtap (Figure 43) shows no systematic decrease due to the TAP analysis: differences are decreased in some areas, but unchanged or even increased in others. An examination of the analysis increments for ETAtap (Figure 44) shows that the analysis increments are larger scale and smaller amplitude than either of the difference fields. The overplotted observation increments are, in most cases, considerably larger than the resulting analysis increments. There are several reasons for this:

- The relative humidity error standard deviations for the observations (15%) are larger than those of the forecast first guess (10%), resulting in a deweighting of the observations.

- The horizontal error correlation function for relative humidity (Figure 31) are large-scale compared to the difference fields.

- The observation increments tend to cancel each other in several areas, where large differences exist between observations that are separated by distances that are much smaller than the error correlation length scales.

- In some areas, there are no observations, such as in the southeast corner of the domain, where there is a local maximum in GLOtap and GLOnotap, but not in ETAtap or ETAnotap.

7.3 Forecast Results

During the MM5 forecast period (12 UTC 7 March 1997 - 12 UTC 8 March 1997) the surface low pressure system associated with the upper-level trough shown in the previous section intensifies and moves northeastward, until it is close to the domain boundary at the final time. The analyzed sea-level pressure field, as derived by DATAGRID from the global NCEP analysis, is shown in Figure 45 for 00 UTC 8 March, or 12 hours into the forecast period. Also shown in that plot are the analyzed positions of the low at the initial and final times, along with the forecast low positions at the final time from the four different MM5 forecasts. The 12-hour forecast sea-level pressure fields from the four separate MM5 forecasts appear quite similar (Figures 46-49). There are, however, some important differences: while the surface low is too deep in all four forecasts, the errors are smaller in the tap than the notap runs, and smaller in the ETA than the GLO runs (Table 4). At the final time, the predicted central pressure agrees better with the analyzed value for all four forecasts, but its position is too far to the northeast, more so in the GLO than the ETA runs. However, at the final time, the surface low is quite close to the boundary of the mesoscale domain, and the solution is likely to be influenced by the interaction with the lateral boundary conditions, making it more difficult to identify the effects of the initial conditions. We therefore focus on the first 12 hours of the forecast in our discussion here.

7.4 Discussion of Results

The results shown in this Section clearly demonstrate the suitability of TAP for intializing the MM5 mesoscale forecast model, when integrated into the MM5 preprocessing suite: forecasts from TAP analyses produced meteorologically reasonable forecasts; comparisons with forecasts from the first guess fields used in the analyses indicate a small, but generally positive analysis impact. For the two first guess fields tested here (a 12-hour ETA model forecast, and a global NCEP analysis), analysis increments and resulting forecast impacts were relatively small. In the case of the humidity analyses, the results indicate that modifications to the error statistics (shorter error correlation length scales, and larger forecast and/or smaller observation errors) might be needed to improve the response of the analysis to the available observations. A possibly even larger analysis impact might be obtained by performing the analysis directly on the MM5 model σ levels. In that configuration, the TAP pre- and postprocessor would need to be modified to include several initialization functions (restoration of hydrostatic balance, and removal of vertically integrated divergent winds) presently performed by the MM5 DATAGRID and INTERP programs.

Table 4: Central pressure (P, in hPa) of the surface low from 00 UTC 7 March (t=0 hr) to 00 UTC 8 March 1997 (t=24 hr)

Name	P at t=0 hr	P at t=12 hr	P at t=24 hr
Analysis	1002	1000	993
GLOnotap		994	993
GLOtap		995	994
ETAnotap		997	992
ETAtap		998	994

A much more pronounced difference in the forecasts is apparent in the 12-hour accumulated precipitation valid at the same time. Comparing the results between the GLOnotap and GLOtap runs (Figures 50 and 51) shows minor differences, most notably somewhat larger accumulations over the Florida panhandle. There are larger differences between the GLOnotap run and the ETAnotap run (Figure 52): the main area of precipitation associated with the surface low is much weaker in the ETAnotap run, consistent with the generally weaker development of the surface low. Another difference is apparent over the Atlantic, where there is an area of precipitation in GLOnotap, but not in ETAnotap. This difference is directly related to the difference in the initial moisture field (Figure 42). One other important difference is that the ETAnotap run has a much larger precipitation maximum to the south, over South Carolina. Interestingly the precipitation field in the ETAtap run (Figure 53) is quite similar to that of ETAnotap everywhere except over the Carolinas, where the values are much smaller, and more in line with the GLOnotap and GLOtap predictions. This difference is consistent with the humidity analysis increments in ETAtap in this area (Figure 44): the high humidity values present in the ETA forecast first guess field were reduced by TAP, based on the available radiosonde observations, to values that more closely agreed with those of the global analysis.

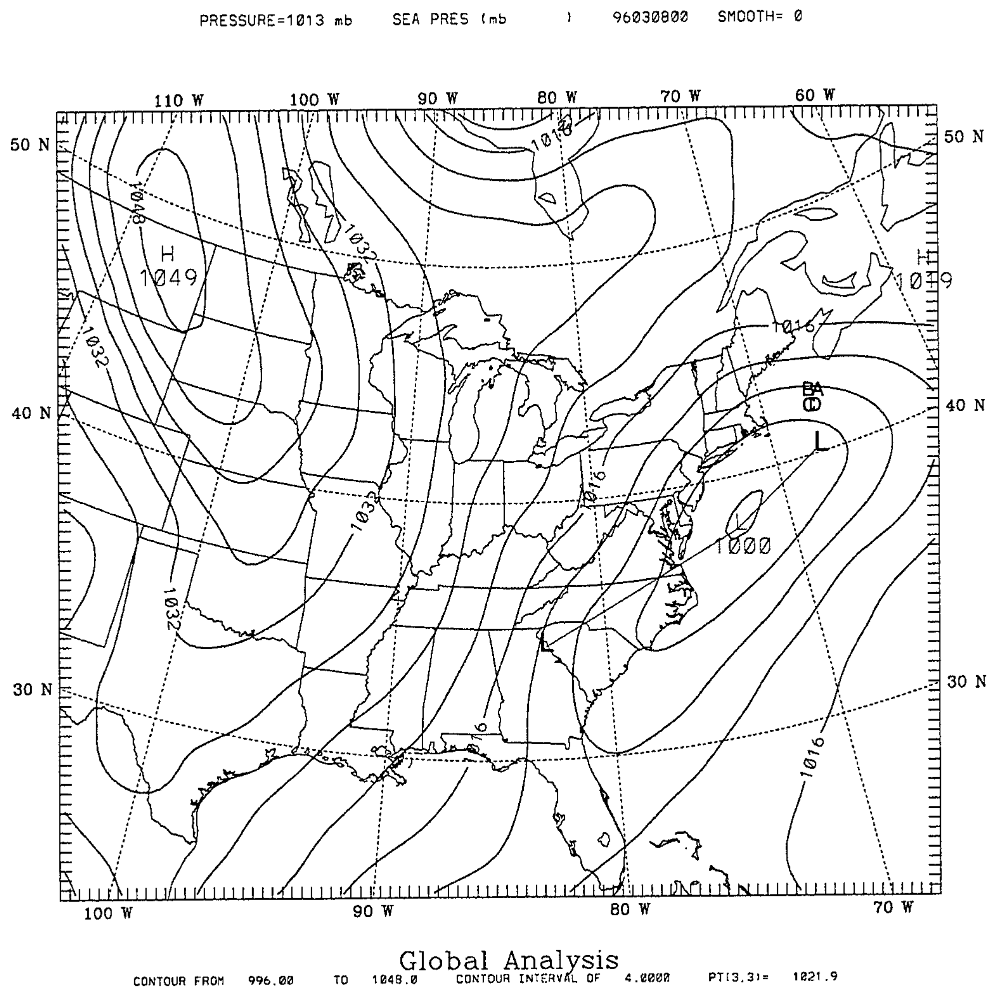


Figure 45: Sea-level pressure analysis generated by DATAGRID from the global NCEP analysis at 00 UTC 8 March 1997. Also shown are the initial and final analyzed positions of the surface low (L, connected by straight lines), and the forecast final low positions for GLOnotap (A), GLOtap (B), ETAnotap (C), and ETAtap (D).

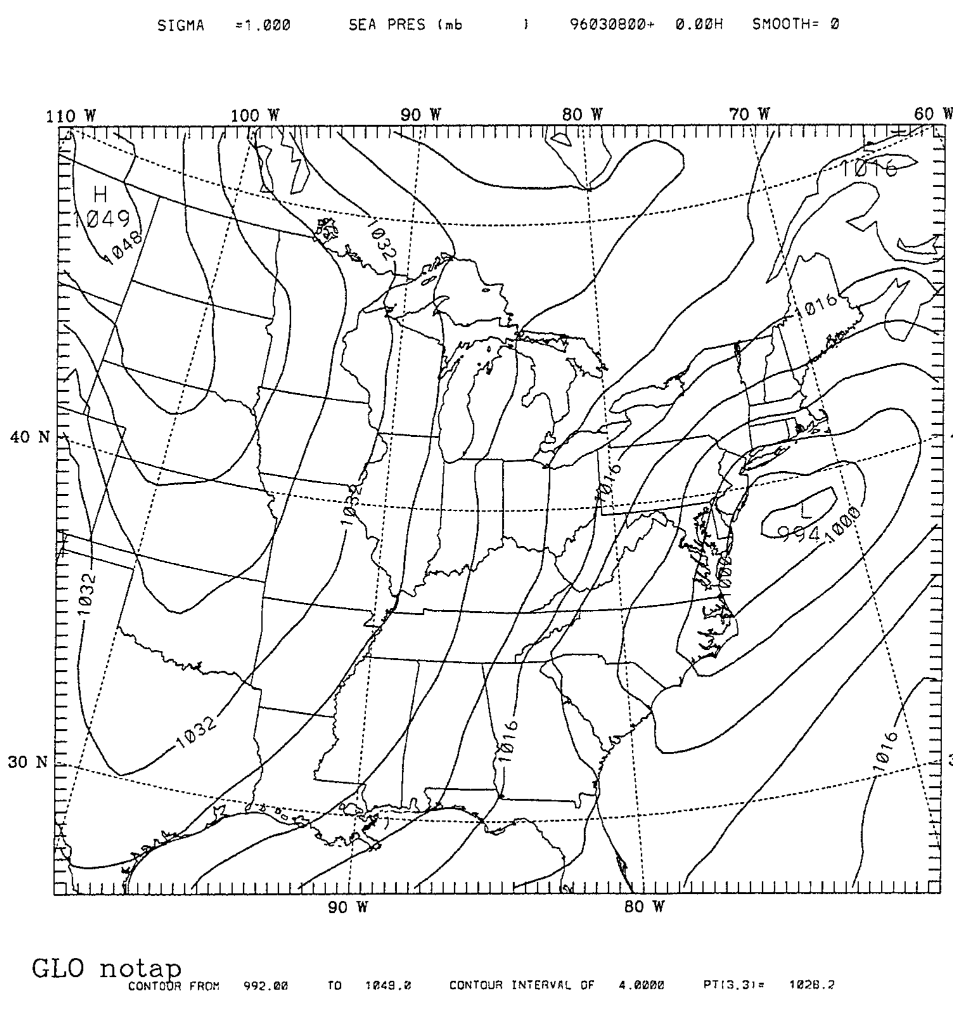


Figure 46: GLOnotap 12-hour forecast of sea-level pressure valid at 00 UTC 8 March 1997.

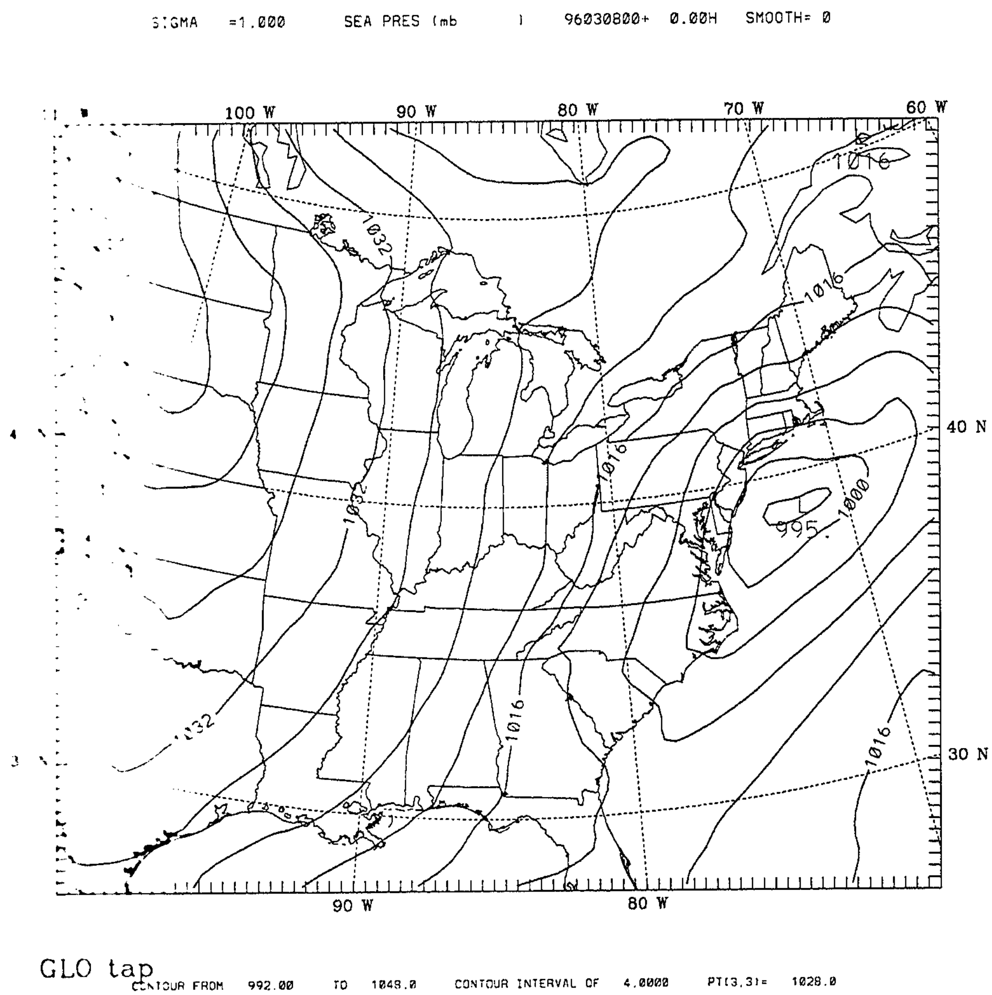
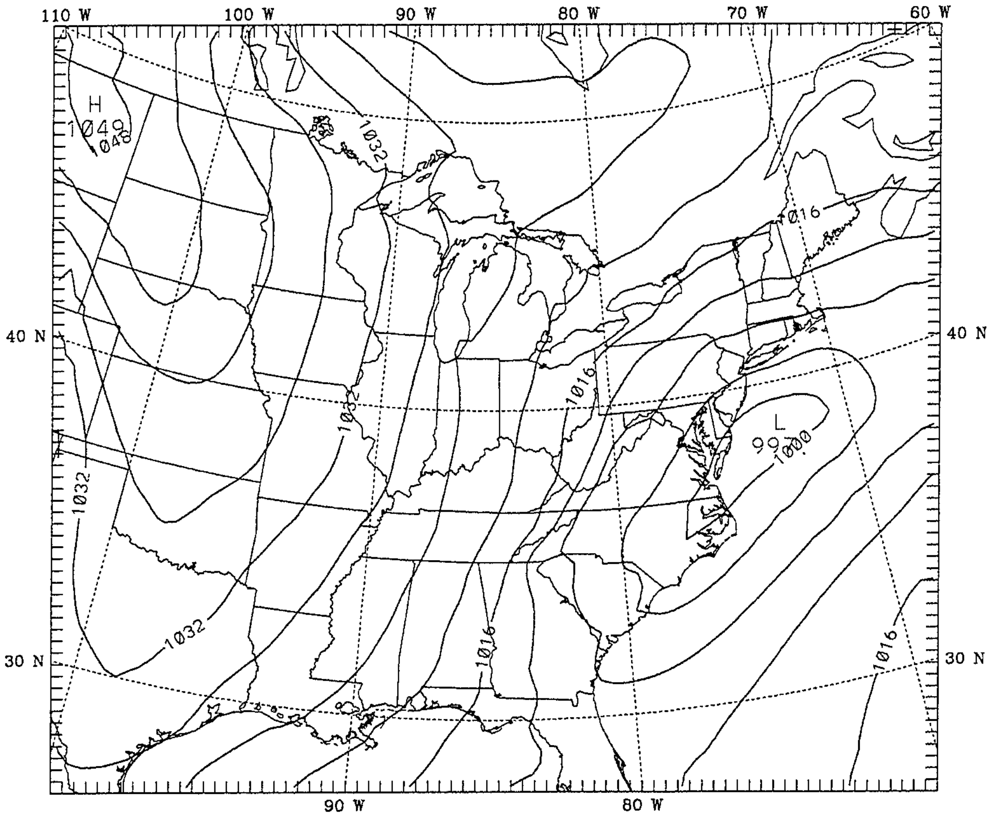


Figure 47: GLOtap 12-hour forecast of sea-level pressure valid at 00 UTC 8 March 1997.

SIGMA =1.000 SEA PRES (mb) 96030800+ 0.00H SMOOTH= 0



ETA notap
CONTOUR FROM 996.00 TO 1048.0 CONTOUR INTERVAL OF 4.0000 PT(3,3)= 1028.2

Figure 48: ETAnotap 12-hour forecast of sea-level pressure valid at 00 UTC 8 March 1997.

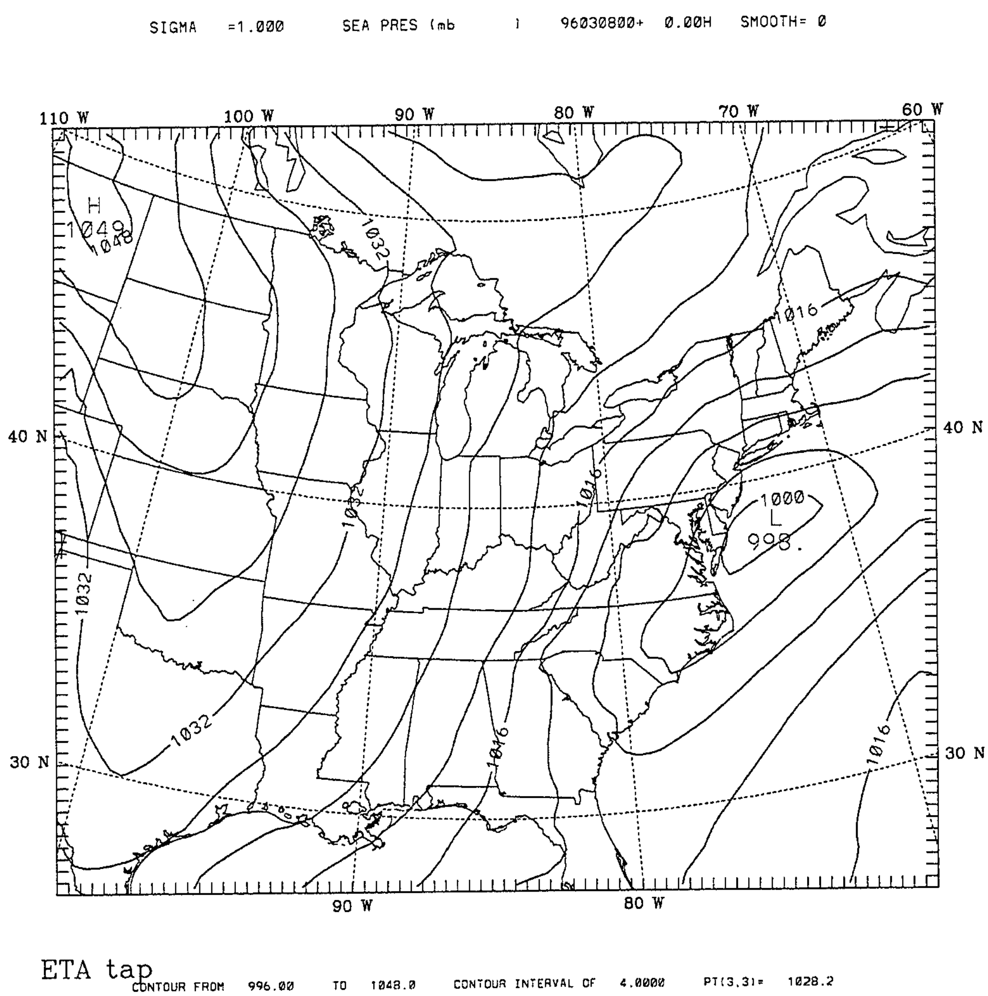


Figure 49: ETAtap 12-hour forecast of sea-level pressure valid at 00 UTC 8 March 1997.

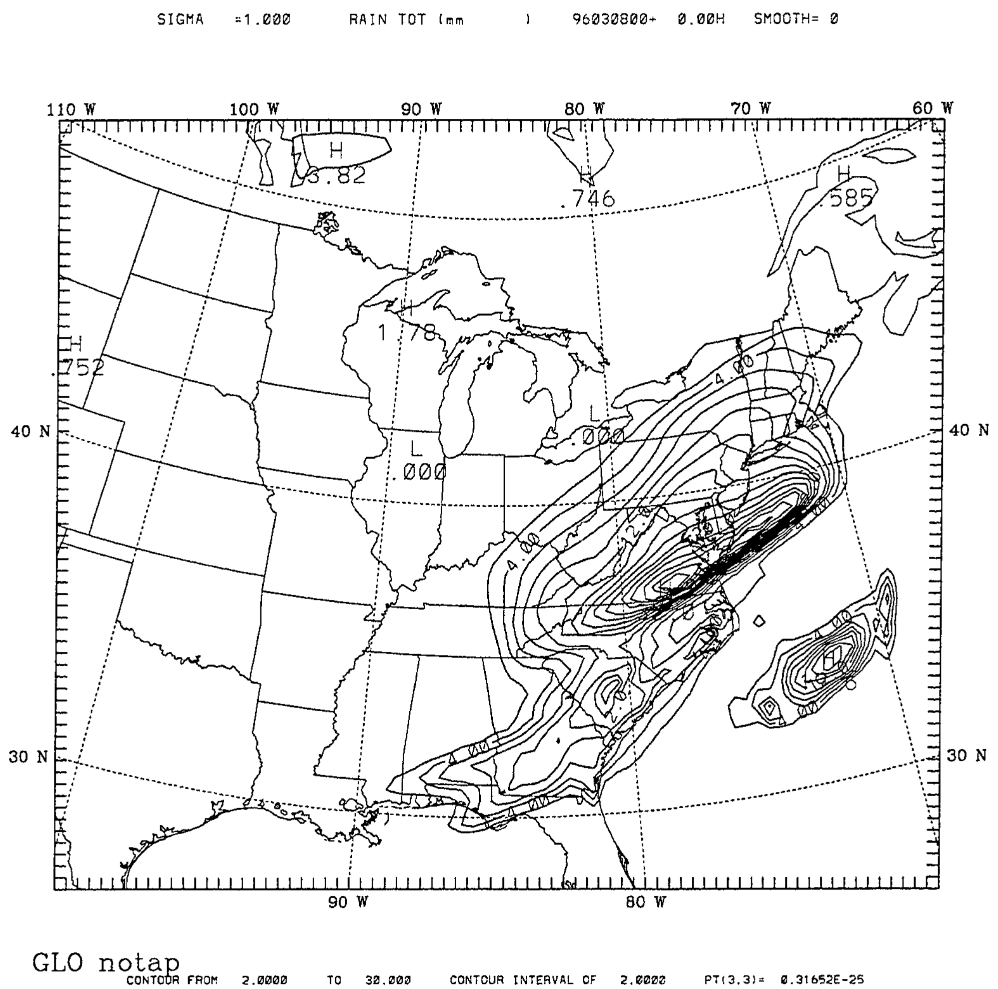


Figure 50: GLOnotap forecast of 12-hourly accumulated precipitation (mm) valid at 00 UTC 8 March 1997.

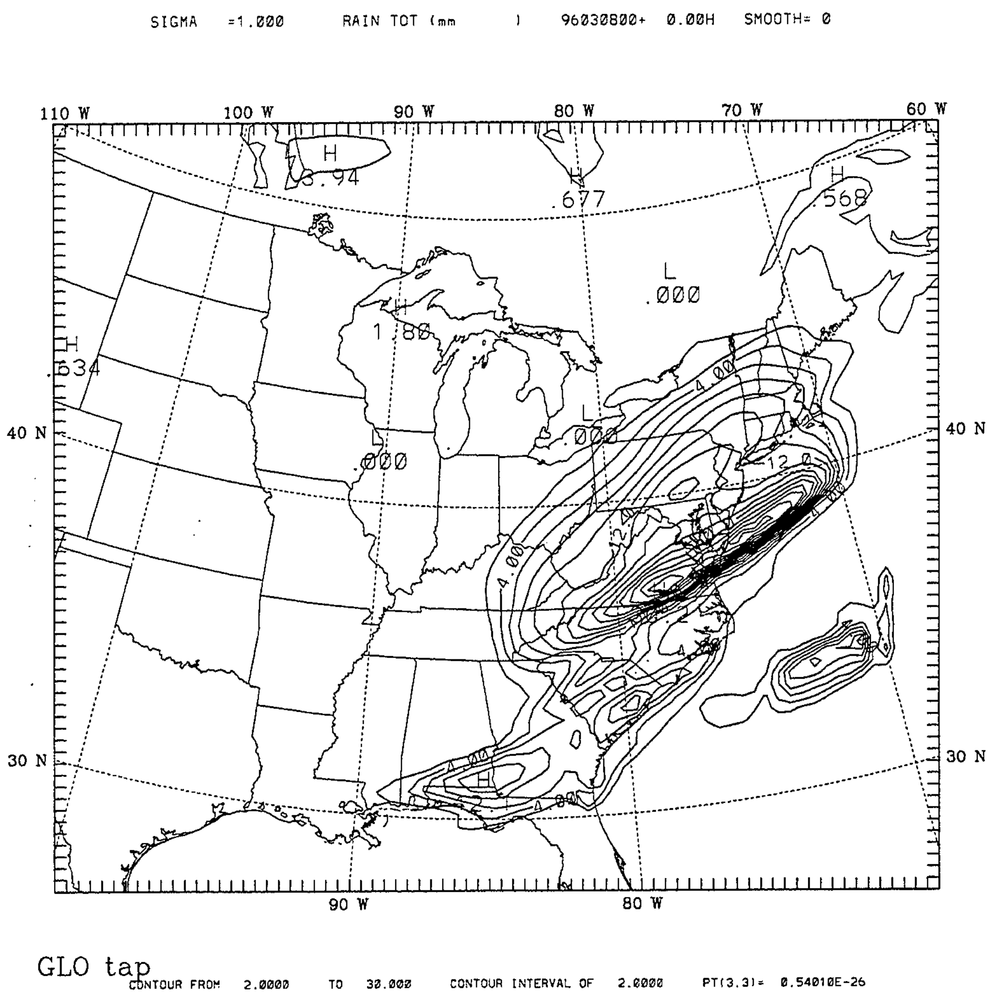


Figure 51: GLOtap forecast of 12-hourly accumulated precipitation (mm) valid at 00 UTC 8 March 1997.

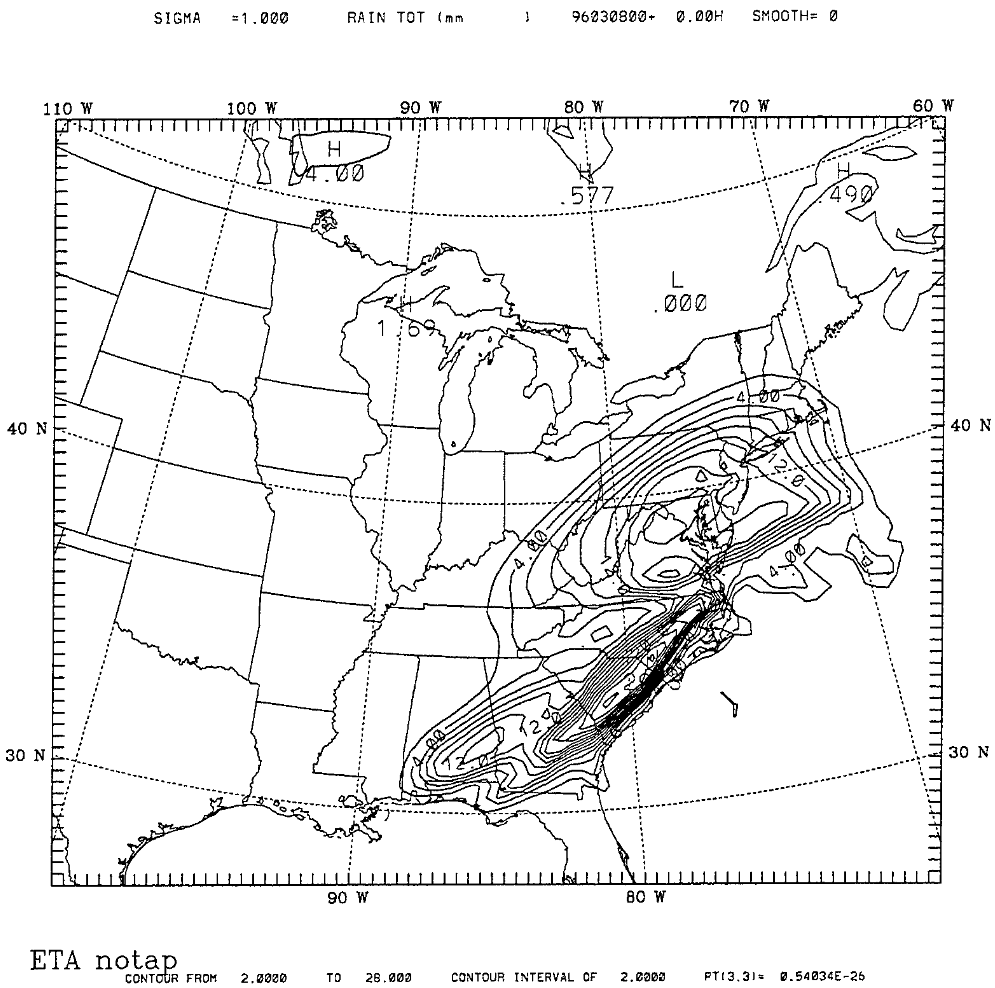


Figure 52: ETAnotap forecast of 12-hourly accumulated precipitation (mm) valid at 00 UTC 8 March 1997.

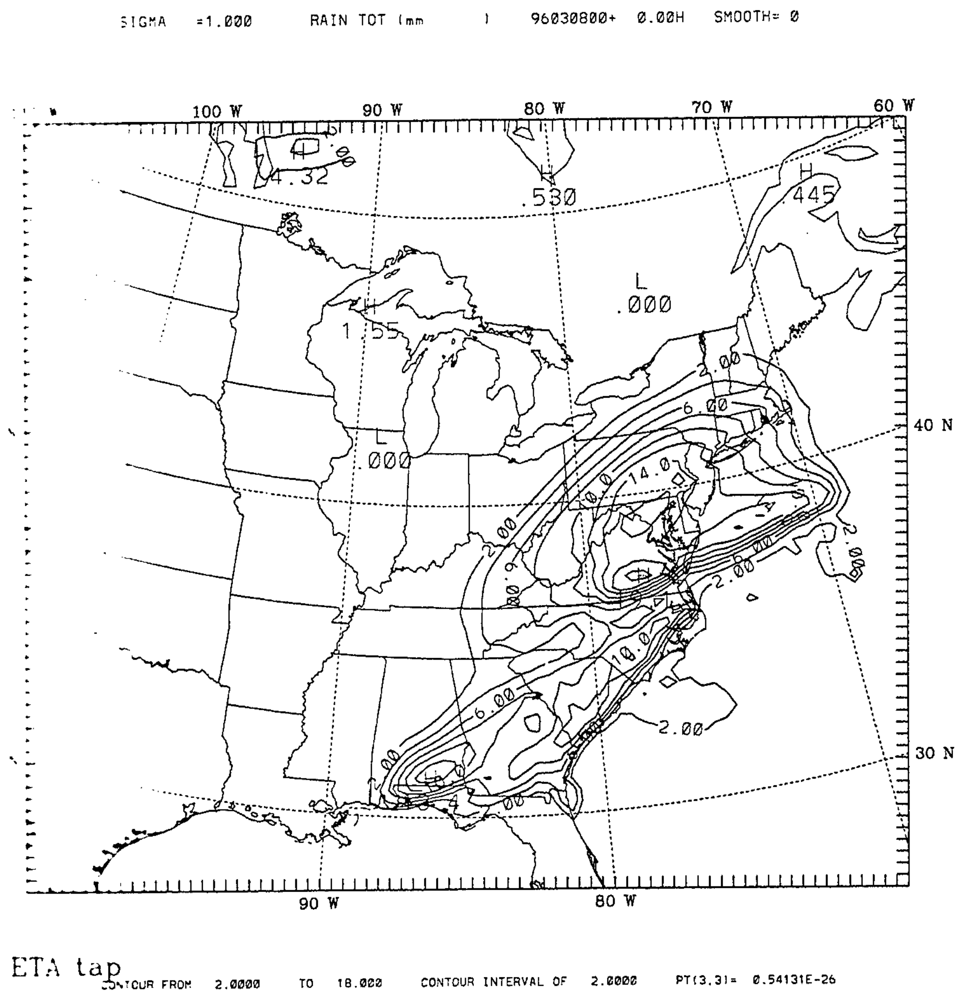


Figure 53: ETAtap forecast of 12-hourly accumulated precipitation (mm) valid at 00 UTC 8 March 1997.

8 System Evaluation at Air Force Weather Agency

The MM5 tests of the TAP prototype were a first step in the evaluation of TAP for the purpose of initializing the MM5 model. A second, more extensive set of tests is planned at AFWA, where TAP is to be evaluated in an operational setting. In preparation for these tests, the TAP software and necessary ancillary data were assembled on a Unix tar tape in a form suitable for installation at AFWA computers. A user's guide including installation instructions was prepared, as well. This guide is included as Appendix C in this report. Because of logistical difficulties at AFWA, the installation could not be accomplished before the end of this contract. However, we tested the software in the same configuration, and using sample observation data files provided by AFWA, to verify its proper functioning.

For the tests planned at AFWA, TAP will be used in essentially the same manner as described in Section 7: TAP combines the first guess field obtained from the DATAGRID output with the available observations to generate a pressure level analysis, which in turn is used as input by the INTERP program. At AFWA, the DATAGRID program is used to reformat and interpolate a short term global forecast from the Navy global forecast model (NOGAPS). In the current operational setting at AFWA, only surface and radiosonde observations are available as input data. For the upper air analysis performed by TAP, only the radiosonde data are used.

References

[AndGM+86] E. Andersson, N. Gustafsson, L. Meuller, and G. Omstedt. Development of meso-scale analysis schemes for nowcasting and very short-range forecasting. Technical Report Promis-Rapporter Nr. 1, Swedish Meteorological and Hydrological Institute, Norrköpping, Sweden, 1986.

[BarM92] P. Bartello and H. L. Mitchell. A continuous three-dimensional model of short-range forecast error covariances. *Tellus*, 44A:217–235, 1992.

[Ben89] S. G. Benjamin. An isentropic mesoα-scale analysis system and its sensitivity to aircraft and surface observations. *Monthly Weather Review*, 117:1586–1603, 1989.

[BenBB+91] S. G. Benjamin, K. A. Brewster, R. Brümmer, B. J. Jewett, T. W. Schlatter, T. L. Smith, and P. R. Stamus. An isentropic three-hourly data assimilation system using ACARS aircraft observations. *Monthly Weather Review*, 119:888–906, 1991.

[BenSM+91] S. G. Benjamin, T. L. Smith, P. A. Miller, D. Kim, T. W. Schlatter, and R. Bleck. Recent improvements in the maps isentropic-sigma data assimilation system. In *Preprints, Ninth Conference on Numerical Weather Prediction*, pages 118–121, Denver, CO, 1991. Amer. Meteorol. Soc.

[Ber79] K. H. Bergman. Multivariate analysis of temperature and winds using optimum interpolation. *Monthly Weather Review*, 107:1423–1444, 1979.

[Bue71] C. E. Buell. Two-point wind correlations on an isobaric surface in a nonhomogeneous non-isotropic atmosphere. *Journal of Applied Meteorology*, 10:1266–1274, 1971.

[Bue72] C. E. Buell. Correlation functions for wind and geopotential on isobaric surfaces. *Journal of Applied Meteorology*, 11:51–59, 1972.

[BurRH+92] John Burgeson, Christopher Redder, Isidore Halberstam, and Ralph Shapiro. Regional-scale analysis and forecasting. Technical Report 92-2227, Phillips Laboratory, Hanscom AFB, MA, 1992.

[Car91] Jean-Marie Carrière. Statistics of horizontal correlation of errors for three-hour prediction fields in isentropic coordinates. In *Preprints, Ninth Conference on Numerical Weather Prediction*, pages 62–65, Denver, Colorado, 14-18 October 1991. American Meteorological Society, Boston, MA.

[Dal91] R. Daley. *Atmospheric Data Analysis.* Cambridge University Press, 1991.

[DeaDB+90] Dennis Deaven, Geoffrey DiMego, Brad Ballish, Lauren Morone, Jean Thiébaux, Paul Julian, and Ralph Peterson. Analysis changes in the global data assimilation system on November 30, 1988. Technical Procedures Bulletin 384, NWS, Silver Springs, MD, 1990.

[DevS94] D. Dèvènyi and T. W. Schlatter. Statistical properties of three-hour prediction "errors" derived from the Mesoscale Analysis and Prediction System. *Mon. Wea. Rev.*, 122(6):1263–1280, 1994.

[DeyM85] C. H. Dey and L. L. Morone. Evolution of the National Meteorological Center global data assimilation system: January 1982-December 1983. *Monthly Weather Review*, 113:304–318, 1985.

[DiM88] G. J. DiMego. The National Meteorological Center regional analysis system. *Monthly Weather Review*, 116:977–1000, 1988.

[DiMMP+92] Geoffrey J. DiMego, Kenneth E. Mitchell, Ralph A. Petersen, James E. Hoke, Joseph P. Gerrity, James T. Tuccillo, Richard L. Wobus, and Hann-Ming H. Juang. Changes to NMC's regional analysis and forecast system. *Weather and Forecasting*, 7:185–198, 1992.

[Eli81] A. Eliassen. Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. In L. Bengtsson, M. Ghil, and E. Källen, editors, *Dynamic Meteorology: Data Assimilation Methods, 1981*, number 36 in Applied Mathematical Sciences, pages 319–330. Springer-Verlag, New York, 1954. Originally published in 1954 as Rapport Nr. 5, Videnskaps-Akademiets Institute for Vaer-Ogklimaforskning, Oslo, Norway.

[Gan63] L. Gandin. *Objective analysis of meteorological fields.* Gidrometeorologicheskoe Isdatel'stvo, Leningrad, 1963. Translated from Russian, Israel Program for Scientific Translations, Jerusalem, 1965. [NTIS N6618047].

[GoeP93] James S. Goerss and Patricia A. Phoebus. The multivariate optimum interpolation analysis of meteorological data at the Fleet Numerical Oceanography Center. FR 7531-92-9413, Navy Research Laboratory, Monterey CA, 1993.

[GreDS94] G. A. Grell, J. Dudhia, and D. R. Stauffer. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). Tech. Note NCAR/TN-398+1A, NCAR, 1994.

[HolL86] A. Hollingsworth and P. Lönnberg. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, 38A:111–136, 1986.

[JulT75] P. R. Julian and H. J. Thiébaux. On some properties of correlation functions used in optimum interpolation schemes. *Monthly Weather Review*, 103:605–616, 1975.

[Lon88] P. Lönnberg. Developments in the ECMWF analysis system. In *Seminar on Data Assimilation Systems and the Use of Satellite Data*, pages 75–120. ECMWF, Reading, U.K., 1988.

[LonH86] P. Lönnberg and A. Hollingsworth. The statistical structure of short-range forecast errors as determined from the radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, 38A:137–161, 1986.

[LonPH86] P. Lönnberg, J. Pailleux, and A. Hollingsworth. The new analysis system. Technical Memorandum 125, ECMWF, Reading, U.K., 1986.

[LonSU92] P. Lönnberg, D. Shaw, and P Undén, editors. *Research Manual 1, ECMWF Data Assimilation, Scientific Documentation.* ECMWF, Reading, U.K., 1992.

[Lor81] A. Lorenc. A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, 109:701–721, 1981.

[McPBK+79] R. D. McPherson, K. H. Bergman, R. E. Kistler, G. E. Rasch, and D. S. Gordon. The NMC operational global data assimilation system. *Monthly Weather Review*, 107:1445–1461, 1979.

[MeuRA+90] Lars Meuller, Lennart Robertson, Erik Andersson, and Nils Gustafsson. Meso-γ scale objective analysis of near surface temperature, humidity and wind and its application in air pollution modeling. Technical Report Promis-Rapporter Nr. 11, RMK Nr. 61, Swedish Meteorological and Hydrological Institute, Norrköpping, Sweden, 1990.

[MilB92] Patricia A. Miller and Stanley G. Benjamin. A system for the hourly assimilation of surface observations in mountainous and flat terrain. *Monthly Weather Review*, 120:2342–2359, 1992.

[NehHS+96] T. Nehrkorn, R. N. Hoffman, J. Sparrow, and M. Yin. Development and testing of Theater Analysis Procedures (TAP): results from year 2. PL-TR 96-2246, Phillips Laboratory, Hanscom Air Force Base, MA., 1996. [NTIS ADA319121].

[NehHY95] T. Nehrkorn, R. N. Hoffman, and M. Yin. Theater Analysis Procedures (TAP) system development: results from year 1. PL-TR 95-2115, Phillips Laboratory, Hanscom Air Force Base, MA., 1995. [NTIS ADA301888].

[Nor88] D. C. Norquist. Alternate forms of humidity information in global data assimilation. *Monthly Weather Review*, 116:452–471, 1988.

[Oor83] A. H. Oort. Global atmospheric circulation statistics. Professional Paper 14, NOAA, Washington, DC, 1983. [NTIS PB8-4-129717].

[ParD92] David F. Parrish and John C. Derber. The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120:1747–1763, 1992.

[RamKS73] Y. Ramanathan, P. Kulkarni, and D. R. Sikka. On a study of winter season wind structure at 500 mb in the Indian region for use in objective analysis of the wind field. *Journal of Applied Meteorology*, 12, 1973.

[RogDD95] E. Rogers, D. G. Deaven, and G. J. DiMego. The regional analysis system for the operational "early" Eta model: original 80-km configuration and recent changes. *Wea. Forec.*, 10:810–825, 1995.

[Rut72] I. Rutherford. Data assimilation by statistical interpolation of forecast error fields. *Journal of the Atmospheric Sciences*, 29:809–815, 1972.

[Sch75] T. Schlatter. Some experiments with a multivariate statistical objective analysis scheme. *Monthly Weather Review*, 103:246–256, 1975.

[SchC91] Thomas W. Schlatter and Jean-Marie Carrière. A correlation model allowing for divergent wind analyses and supporting three-hour assimilation of mesoscale data. In *Preprints, Ninth Conference on Numerical Weather Prediction*, pages 58–61, Denver, Colorado, 1991. American Meteorological Society, Boston, MA.

[ShaLH+87] D. B. Shaw, P. Lönnberg, A. Hollingsworth, and P. Undén. The 1984/85 revisions of the ECMWF assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 113:533–566, 1987.

[Thi85] H. J. Thiébaux. On approximations to geopotential and wind-field correlation structures. *Tellus*, 37A:126–131, 1985.

[Thi75] H. Jean Thiébaux. Experiments with correlation representations for objective analysis. *Monthly Weather Review*, 103:617–627, 1975.

[Thi76] H. Jean Thiébaux. Anisotropic correlation functions for objective analysis. *Monthly Weather Review*, 104:994–1002, 1976.

[Thi77] H. Jean Thiébaux. Extending estimation accuracy with anisotropic interpolation. *Monthly Weather Review*, 105:691–699, 1977.

[Und89] Per Undén. Tropical data assimilation and analysis of divergence. *Monthly Weather Review*, 117:2495–2517, 1989.

A Development of a Small-Scale, Relocatable Optimum Interpolation Data Analysis System

Conference paper contained in the Preprints of the 11th Conference on Numerical Weather Prediction of the American Meteorological Society, held in Norfolk, Virginia, August 1996.

DEVELOPMENT OF A SMALL-SCALE, RELOCATABLE OPTIMUM INTERPOLATION DATA ANALYSIS SYSTEM

Thomas Nehrkorn * and Ross N. Hoffman
Atmospheric and Environmental Research, Inc.
Cambridge Massachusetts

1 INTRODUCTION

A prototype optimum interpolation (OI) analysis system is being developed to provide detailed atmospheric analyses under a variety of conditions. The analysis system is designed to be part of a meteorological workstation with its own system for receipt, storage, and display of meteorological data. Its output can be used for display or initialization of locally run mesoscale models. The system is designed to be flexible to adapt to different user requirements: it can generate two or three-dimensional, univariate or multivariate (mass-wind) analyses; it can produce analyses on regular grids using a number of different map projections, or on arbitrarily spaced analysis points; it is able to utilize different background (or first guess) fields, ranging from mesoscale or large-scale model forecasts to climatological background fields; it can be configured to operate over any region of the globe; it can accept a wide variety of observations. Aspects of the system design are reviewed in the next section, results from a preliminary prototype version of the analysis system are shown in section 3, and future plans are discussed in section 4.

2 SYSTEM DESIGN

The analysis system consists of a preprocessor component, analysis component, and postprocessor component.

The preprocessor is responsible for ingest and reformatting of the background, observation, and auxiliary data. This includes preliminary quality control checks of the observations: a check on the magnitude of the difference from the background, and a median filter "buddy check" for some densely spaced data (particularly satellite-derived observations). If needed, the background fields are also interpolated to the analysis grid as part of the preprocessor, and input variables are transformed to those used within the analysis (e.g., dew point to specific humidity, grid-relative wind components to east-west components).

The postprocessor performs similar functions, in reverse: variable conversions from analysis to display variables, regridding from analysis to display grids, reformatting from internal to database formats, and storage of the analysis output in the workstation database.

In the analysis component the background and observations are combined using the standard OI formulation (Lorenc, 1981). Analysis values (A) at each of the grid points are obtained by a linear combination of the first guess (P) and a weighted sum of the surrounding observation increments ($O - P$):

$$A = P + \boldsymbol{w}^T(\boldsymbol{O} - \boldsymbol{P}),$$

where the weights $\boldsymbol{w}$ are determined from the normal equation

$$\mathbf{M}\boldsymbol{w} = \boldsymbol{h}.$$

Here $\mathbf{M}$ is the symmetric positive definite matrix with elements given by

$$m_{ij} = \langle \pi_i \pi_j \rangle + \epsilon_i^o \langle \beta_i \beta_j \rangle \epsilon_j^o,$$

and the elements of $\boldsymbol{h}$ are given by

$$h_i = \langle \pi_i \pi_k \rangle.$$

In these equations, the terms $\langle \pi\pi \rangle$ are the backgound error correlations and the terms $\langle \beta\beta \rangle$ are the observational error correlations, which are multiplied by the ratio (ϵ^o) of observation to background error standard deviations. Correlations between background and observational errors are assumed to be zero.

At present, the analysis equations are solved using the volume method, in which a single matrix equation is inverted for all analysis grid points within a

*Corresponding author address: Thomas Nehrkorn, AER, Inc., 840 Memorial Drive, Cambridge, MA 02139. e-mail: nehrkorn@aer.com

specified volume. An alternative method will be implemented, in which the normal equations are solved point by point using a small subset of observations around each analysis point. This method has the advantage of faster execution times and a straightforward implementation of the OI buddy check procedure, but the resulting analyses may be noisy because of differences in data selection of neighboring gridpoints.

The background and observation error covariances are computed from the corresponding standard deviations and correlations. Correlations are modeled as separable function, i.e. as the product of vertical and horizontal (along pressure surfaces) correlations. To accomodate the desired flexibility of the system with respect to geographic location, the resolution of the input and output fields, and the type of background and observation data, the specification of the required error statistics is separated from the rest of the system design. Observation and background error standard deviations are stored in tables. Horizontal and vertical correlation functions for background and observation errors are also stored in tabular form, with an option to generate them from functional fits to empirical data. Correlations involving wind components are computed using the natural coordinate system of longitudinal and transverse wind components, after Daley (1991).

An example plot of horizontal height-height error correlations is shown in Fig. 1 for the second-order autoregressive function (SOAR) (Goerss and Pheobus, 1993) used for global forecast model background fields. The curves labeled "f" and "g" refer to the (rescaled) first and second derivative of the correlation function (labeled "z-z"). The autocorrelations of the transverse and longitudinal wind components are computed as linear combinations of these two functions.

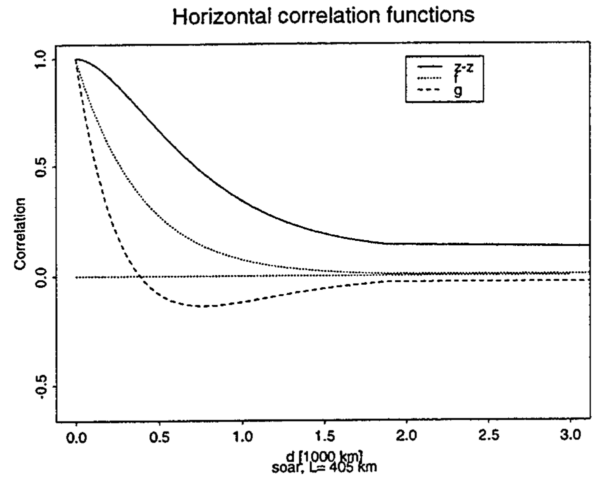


Figure 1: Horizontal height error correlation functions for global model backgrounds (see text).

3 PRELIMINARY RESULTS

The prototype system has been tested on historical data sets. Fig. 2 shows the 700 hPa wind field from a height-wind analysis for one case, hurricane Opal at 00 UTC 5 October 1995, using the ETA model 12-hour forecast as a first guess. The corresponding analysis increments (and observation residuals) are shown in Fig. 3, indicating that the analysis weakened the circulation and moved it to the east relative to the first guess. Of particular interest are the negative height increments over Tallahassee and Tampa, which induce a cyclonic circulation in the wind increments over northeast Florida. This feature is completely absent if height observations are not used in the analysis (Fig. 4): the double-vortex structure of the wind increments in the multivariate analysis is replaced by a single large anticyclonic circulation in the winds-only case.

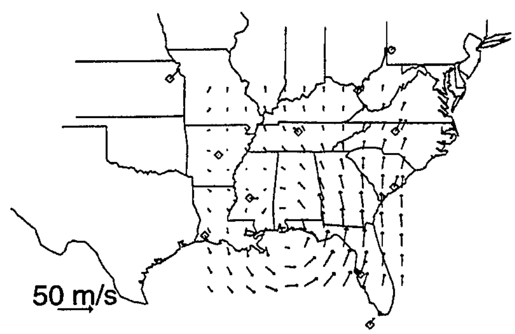


Figure 2: Wind field from a height-wind analysis at 700 hPa for hurricane Opal, using the 12-hour ETA model forecast as a first guess. Raobs used in the analysis are plotted as vectors originating from diamond symbols.

4 FUTURE PLANS

Capabilities that will be added to the existing prototype include:

- an OI "buddy check" procedure in which each observation is compared to an analyzed value

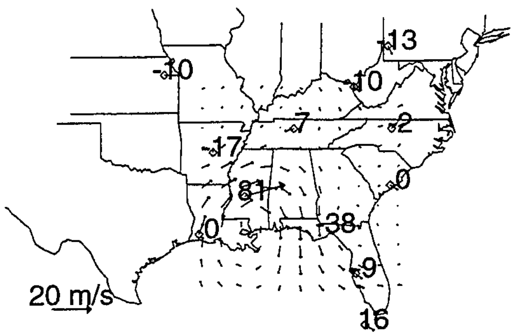


Figure 3: Wind analysis increments at 700 hPa for hurricane Opal, from a multivariate height-wind analysis. Also shown are observation increments for height (values in m) and winds (vectors).

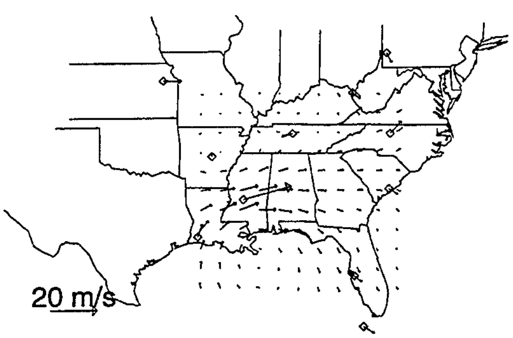


Figure 4: Wind increments at 700 hPa for hurricane Opal, from a winds-only analysis.

derived from surrounding observations

- a point-by-point method for data selection based on stepwise regression (Jennrich, 1977), in which observations are added (or deleted) based on the correlations with the analysis grid points and the already selected observations.

In addition, key system components (data selection, computations of correlations, normal equations solution) will be refined for optimizing performance and execution speed.

References

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press.

Goerss, James S., and Phoebus, Patricia A., 1993: *The Multivariate Optimum Interpolation Analysis of Meteorological Data at the Fleet Numerical Oceanography Center.* FR 7531-92-9413. Navy Research Laboratory, Monterey CA.

Jennrich, Robert I., 1977: Stepwise Regression. *Chap. 4 of:* Enslein, Kurt, Ralston, Anthony, and Wilf, Herbert S. (eds), *Statistical Methods for Digital Computers.* Mathematical Methods for Digital Computers, vol. III. New York: Wiley.

Lorenc, A., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, **109**, 701–721.

Nehrkorn, T., Hoffman, R. N., and Yin, M., 1995: *Theater Analysis Procedures (TAP) system development: results from year 1.* PL-TR 95-2115. Phillips Laboratory, Hanscom Air Force Base, MA. [NTIS ADA301888].

5 Acknowledgements

We have benefitted from helpful discussions with Don Norquist. Data was obtained from the National Center for Environmental Prediction FTP server and the Air Force Interactive Meteorological System (AIMS) operated jointly by AER and the USAF Phillips Laboratory. This research is supported by USAF contract F19628–94–C–0027.

B Early Prototype User's Manual for Testing at the Combat Weather Facility

User's Manual prepared for the real-data tests of TAP by Air Weather Service personnel at the Combat Weather Facility in Hurlburt Field, Florida.

Contents

List of Figures

List of Tables

1 Introduction

The Theater Analysis Procedures (TAP) are being developed to provide a meteorological analysis capability that will reside on a computer workstation. The prototype system is being tested early on in the development cycle by Air Weather Service (AWS) personnel at the Combat Weather Facility (CWF). The testing serves the dual purpose of familiarizing AWS with TAP capabilities, and providing feedback to the TAP developers about strengths and weaknesses of the system.

2 Scope

2.1 Identification

This document is the User's Guide for the early prototype testing of the TAP at the Combat Weather Facility (CWF). It applies to the interim version of TAP after 2 of 3 years of the basic TAP development effort. Refer to section 2.3 for a reader's guide.

2.2 System overview

2.2.1 Objectives of the TAP Project

The TAP project primary objective is to develop robust analysis procedures to support the tactical user. These analysis procedures provide stable meteorological products for end users.

TAP is modular, and capable of utilizing a variety of background and data sources. This capability allows TAP to adapt to different theater meteorological support systems (TMSSs), run on different platforms and satisfy different user requirements. TAP is configurable to a range of requirements, from first-in stand-alone capability to full Theater Weather Central (TWC) support.

2.2.2 Major Functions of TAP

The function of TAP is to use the optimal interpolation technique to combine background (i.e. *a priori*) information with observations of diverse type, quality, and density to produce analyses of meteorological fields. The TAP analysis configurations are optimized to initialize NWP models and to provide input for TDAs.

2.2.3 Performance Issues

TAP is required execute in minutes on a modest workstation. TAP is required be robust. Note that the early prototype version will not satisfy all the performance requirements because less efficient, but more general and robust methods are used in parts of the system. Performance bottlenecks will be identified and replaced by faster code in the remainder of the TAP development.

2.2.4 Management and Technical Constraints

TAP is state of the art, but not experimental. TAP is designed to work every time. In measuring the quality of TAP, robustness of the system is weighted heavily.

2.3 Document overview

This document gives a brief overall description of the TAP project, provides instructions for installing and running the TAP system, and describes the test cases provided for this early prototype test at the Combat Weather Facility (CWF).

At a minimum, test personnel should read Sections 5.1 and 5.2 for basic operation of the system. For interpretation and evaluation of the test results, test personnel may refer to Section 6. For troubleshooting, test and support personnel may refer to Section 5.5.

3 Referenced documents

For more detailed TAP documentation, see also the System and Interface requirements specifications and design documents (document `srs`, document `sdd`, document `irs`, and document `idd`). The plan for testing at the CWF is described in a memorandum by the Phillips Laboratory dated 15 March 1996 (revised 30 April 1996). This memorandum contains a decription of the test objectives and schedule, which is not repeated here. Detailed test evaluation criteria are provided in the form of a questionnaire which is distributed to all test personnel.

4 Installing TAP

4.1 Required system software

The following system software is assumed to be preinstalled on the host system:

- Solaris 2.4 operating system
- Sun Fortran 77 compiler
- Sun C compiler.

4.2 Installing supporting software

4.2.1 Splus

The TAP system makes extensive use of the Splus software package. It is a proprietary software package for interactive data analysis and display that is provided with its own storage medium and installation instructions. The location of the Splus software (both the location of the Splus Home directory and the directory of the executable Splus script) must be recorded for later use in the TAP installation.

4.2.2 Netscape server and browser

The graphical user interface of TAP makes use of the Netscape software packages for running an http (HyperText Transfer Protocol) server and browser. It is a proprietary software package that is provided with its own storage medium and installation instructions. For purposes of the TAP demonstration, a copy of the needed software has been placed on the tar tape along with the other binaries (see below). The browser is automatically installed by performing the tar command described below. Before running the installation script for the server (`server/httpd/install/ns-setup`), the directory containing the browser (or a link to it) must be included in the `PATH`. The location of the Netscape browser executable must be recorded for later use in the TAP installation. The Netscape server must be configured to enable execution of CGI (Common Gateway Interface) scripts by editing of the files `config/magnus.conf` and `config/obj.conf` files (either manually or with the supplied Server Manager). The host name of the server machine, and the http address of the directory containing the cgi scripts, must be recorded for later use in the installation. Some system files may need to be edited for proper operation of the server.

4.2.3 Others

All other software packages have been assembled onto a tar tape and can be directly restored from tape to disk. For the basic TAP system, they are: GNU Make and related utilities; Ghostview, xv, and related utilities; perl; LAPACK and BLAS source code and binary files. To support the generation of postscript documents from the LaTeX source files, the LaTeX package is also included in the basic package of supporting software. The location of the binaries must be noted for later use in the TAP installation.

The supporting software, including Netscape, are installed by a simple tar command. For example, if the software is to be stored in directory `/users/cwf/tap.demo/`, then the required Unix commands are

```
cd /users/cwf/
mkdir tap.demo
cd tap.demo
tar -xvf /dev/rmt1 |& tee tar.out1
```

where `/dev/rmt1` denotes the tape drive on which the tar tape is mounted. Upon completion of this command, several new directories are created, which contain all the supporting software. A listing of all the files is contained in file `tar.out1`.

Not included in the basic package are utilities and programs that are only needed for the software development environment: RCS for version control, Gnu Emacs for editing of files and system-level interactive use of the TAP software.

4.3 Installing the TAP software

The TAP software and ancillary datasets are all stored together under the `TAPHOME` directory. For the purpose of the early prototype test, since both the machine used during TAP

development and for the early prototype are Sun workstations, binary versions of the TAP code and data sets can be directly installed on the host system. The entire directory tree structure is stored on a tar tape, and installation only requires transfer from the tar tape onto the appropriate directory on the disk of the host system. For example, if the desired disk location is `/users/cwf/tap.demo/`, then the required Unix commands are

```
cd /users/cwf/tap.demo
tar -xvf /dev/rmt1 TAPHOME |& tee tar.out2
```

where again `/dev/rmt1` denotes the tape drive on which the tar tape is mounted. Upon completion of this command, a new directory `TAPHOME` is created, which contains all the TAP software and ancillary data. A listing of all the files is contained in file `tar.out2`.

After transfer of the data to disk, some editing of files is required to change path names of certain system and supporting software. A cshell script (`install.csh`) in the `TAPHOME` directory is provided for this purpose. The script creates and executes a file of editor commands (for the UNIX `sed` editor), based on user responses to queries for locations of the various software packages and system binaries. Default values are provided for most of these. Refer to file `TAPHOME/README` for up-to-date instructions.

As a final step in the installation, files with filename extension `cgi` in directory `TAPHOME/html` must be copied to the location designated for CGI scripts during the Netscape server installation.

4.4 Adding or removing TAP users

TAP is set up to support multiple users. Separate disk areas are provided for each user for storage of output and intermediate files. Adding a user, (for example: `bob`) simply requires the addition of subdirectories by the commands:

```
mkdir $TAPHOME/users/bob ; mkdir $TAPHOME/users/bob/.Data
```

Conversely, removing this user is accomplished by

```
"rm" -fr $TAPHOME/users/bob
```

For the usual, single-user mode of operation, the user name `tap` is preinstalled on the system.

5 Running TAP

Two ways of running the TAP system are provided for the early prototype testing at CWF: cshell scripts which are invoked by the user from the system prompt after login, and a graphical user interface which requires running a Web server (Netscape server) on the host system and a Web browser (such as Mosaic or Netscape) by the user. It is anticipated that the graphical user interface will be the preferred way of running TAP, but the cshell scripts are provided as a backup and alternative for users with more Unix experience.

5.1 Required initializations for the Unix session

Before invoking TAP through either the graphical or cshell user interface, the environment variable TAPHOME must be set to the location of the tap software, e.g. in the above example it would be

```
setenv TAPHOME /users/cwf/tap.demo/TAPHOME
```

Further environment variables and command aliases are then set by issuing the command

```
source $TAPHOME/setenv.csh
```

Note to system administrator: Both of these commands could be included in the user's `.cshrc` file, so that they will be executed automatically upon login.

A simple help command (`tap.help`) is provided that lists all available tap commands and the user's TAP configuration. It is described in more detail in Section 5.3.

5.2 The graphical user interface

The graphical user interface is invoked by the commmand `tap.web &`. It does not matter from what directory this command is invoked. This will start the Web browser and open the TAP `main-menu` page (see Figure 1 for an approximate rendition of the display as seen on a computer terminal).

When displayed on a computer terminal, "hyperlinks" to other TAP pages are specially marked text (usually underlined); moving the cursor to such a hyperlink and clicking the left mouse button will display the referenced page. The main-menu page can be displayed again by clicking on the main-menu links in the other TAP pages, or by using the "Back" command of the Web browser (using the "Go" pull-down menu in the case of Netscape version 2). Both the TAP `overview` and `documentation` pages provide background information on TAP and require no further explanation here. The `View gif images of NWS products` page is a link to a directory containing gif images in several subdirectories. Clicking on this link will display a listing of that directory. From a directory listing, one can view listings of subdirectories or images of gif files by clicking on the approriate directory or file name. The date and time and type of plot are apparent from the directory and file names. The remaining pages (`execute`, `plot`, `status check`, and `remove`) are hypertext "forms", containing a number of input fields that are either entered from the keyboard, or selected by mouse clicks on specially marked "button" icons. Once filled out with all the required inputs, they are submitted by clicking on a button at the bottom of the page which is labeled with the action to be performed. Submitting the form initiates processing, and causes a new page to be displayed with output generated from the process. These pages are described in more detail below.

TAP

Theater Analysis Procedures

AER, Inc. USAF Phillips Laboratory

Caution: TAP is under construction.
Your help is appreciated. Please send e-mail to the address below.

TAP main menu

- TAP overview
- TAP documentation
- Check status of TAP analysis jobs
- Execute TAP
- Plot a TAP map
- View gif images of NWS products
- Remove TAP output from user directory

AER, Inc., Cambridge, MA
rhoffman@aer.com

This is Theater Analysis Procedures (TAP) html documentation.

The TAP project (AER P584) is funded by the USAF (F19628-94-C-0027).

AER, Inc. intends to retain patent rights to certain aspects of the TAP algorithms under FAR 52.227-11.

Revision control: `$Id: main.menu.html,v 1.4 1996/04/29 20:47:21 trn Exp trn $`

Figure 1: TAP Main menu page

5.2.1 Status check page

Before executing a TAP analysis, one must make sure there are no other TAP analysis jobs running. Before plotting results from a TAP analysis, one needs to know the name given to the analysis run, and make sure it has completed. The Status check page (see Figure 2) provides a convenient way of accomplishing this.

```
Check on TAP cases

   * Analysis Name:
 - There are 2 Options:
 --(1): list names and brief description of all cases: Leave Analysis Name blank
 --(2): provide a full list for a single case: Analysis name must be specified

   * Analysis owner:
```

Figure 2: Required inputs for the TAP status check menu page. Analysis Name and Owner are text fields to be entered.

Submitting this form will launch a unix command to examine the TAP user directory for analysis output files. The output will be displayed to the screen.

Using an empty (blank) field for Analysis Name will produce a list of all analysis jobs and indicate whether they have finished or are still running. Also listed for each analysis job is its timestamp file, which contains information on the start and end time, and the levels and variable codes used for the analysis. Specifying an analysis name will display, in addition, the size of all data files for that analysis, and the contents of the log file. This display is useful for debugging purposes, but should not usually be required. Interpretation of the log file should be referred to the software programmer.

The TAP user name (also referred to as "analysis owner") identifies a disk space unique to that user. The only valid user names are those for which directories have been established at the time of installation. The default value ("tap") is the name reserved for the usual, single-user mode of operation of TAP. The TAP user name has no connection to the Unix login or user name.

5.2.2 Execute page

This link is selected to execute TAP and produce an analysis. The required user inputs are shown in Figure 3.

Submitting this form will launch an analysis job. The input specifications will be listed to the screen immediately. Results may be plotted upon completion of the job (see Section 5.2.3). The status of TAP analysis jobs may be checked using the Status check page (see 5.2.1) to make sure no other jobs are running, and to check on the status of the current job.

Warning: No other analysis jobs should be running at this time in the TAP user directory.

Execute TAP

Set the following parameters to specify a TAP analysis.

* Data source:
  - March 95, N.E. US (12 UTC 6 March 1995)
  - March 95, S.E. US (12 UTC 6 March 1995)
  - Opal, N.E. US (00 UTC 5 October 1995)
  - Opal, S.E. US (00 UTC 5 October 1995)
* Domain:
  - Large analysis domain using 150 km resolution
  - Small analysis domain in center using 75 km resolution
  - Column at domain center (with 75 km effective resolution)
* Analysis scheme:
  - Height/wind analysis (upper air). codes= (7,33,34)
  - Height analysis (upper air). code= 7
  - Wind analysis (upper air). codes= (33,34)
  - Temperature analysis (upper air). code= 11
  - RH analysis (upper air). code= 52
  - Surface Temperature analysis (2d). code= 11
* Background type:
  - Climatology
  - 12-hour ETA model forecast
* Analysis name:

* Analysis owner:

Figure 3: Required inputs for the TAP execute menu page. Analysis name and owner are text fields to be entered, all others are choices that are selected by clicking a button icon next to the desired value.

For purposes of the prototype tests, TAP is restricted to running a number of canned cases over prespecified regions and analysis domains. The Data Source selection identifies the case (date and time) and geographic region, and the analysis domain one of the three available analysis grid options. The cases and grid options are described in more detail in Section 6. The analysis scheme selects the type of analysis to be produced. All upper air analyses are at a preselected set of pressure levels (1000 mb; 925 mb; 850 mb; 700mb; 500 mb; 400 mb; 300mb; 250 mb; 200 mb; 150 mb; 100 mb).

The name of the analysis must be a legal UNIX file name and a legal Splus variable name. This is always guaranteed if it begins with a letter and contains only alphanumeric characters. It is important to make a note of the name of the analysis, because this name is needed to generate plots.

Warning: Reusing a name will overwrite a pre-existing analysis.

The Analysis owner must be a valid TAP user name (see Section 5.2.1).

5.2.3 Plot page

Upon completion of the TAP job, results may be plotted using the plot page shown in Figure 4. Plots will be plots of one or more variables at the analysis grid points, at one of the analysis levels, over a map background of the analysis domain. Values of observations used in the analysis may optionally be overlayed over the plot.

Submitting this form will launch a job that will create graphical output (in postscript format) in the specified file (default: `tap.ps`) in the user's directory. A listing of specified plot parameters is echoed to the screen. Upon completion of the plot job, a postscript viewer (ghostview) is invoked to display the graphical output to the screen. The ghostview window contains pull down menus for printing a hard-copy of the output, saving it to a file with a different name, magnification/reduction, and numerous other display options.

The name of the analysis must be specified. It must be the name of a previously completed analysis job (see Sections 5.2.2 and 5.2.1).

The analysis owner must be a valid TAP user name, and it must be the owner of the analysis to be plotted.

For variables other than wind vectors, contour plots are drawn if the analysis name is an analysis over the small or large domain. If the analysis is for a single column, the value is printed at the analysis point location. Other display options (e.g., for vertical profiles or cross sections) may be added at a later date. The code numbers displayed with the choice of variables refer to the code numbers assigned to each variable within TAP. These code numbers are also displayed on the title of the plots. Selecting either a variable or a level that is not present in an analysis will produce a message (in the postscript output file and the display generated from it) listing the available levels and/or variables.

The plot type options identify what type of values are displayed: either analyzed values, background values (interpolated to the analysis locations from either the forecast or climatology first guess), or increments (the difference between the first guess and the analysis values). For increment plots, observation plots will be those of observation increments (i..e., the difference between the observed value, and the background value interpolated to the observation location).

Plot a TAP map

Set the following parameters to specify a TAP map.

* Analysis name:

* Analysis owner:

* Variable:
  - Geopotential height (Z) [gpm]. Code=7
  - Wind vectors [m/s]. Code=33 34
  - Geopotential height (Z) [gpm] plus Wind vectors (u,v) [m/s]. Code=7 33 34
  - Temperature (T) [K]. Code=11
  - Relative humidity (RH) [%]. Code=52
* Level:

Surface; 1000 mb; 925 mb; 850 mb; 700mb; 500 mb; 400 mb; 300mb; 250 mb; 200 mb; 150 mb; 100 mb

* Plot type:

Analysis; Background; Increments

* Overlay observation values:

True; False

* Postscript file name:

Figure 4: Required inputs for the TAP plot menu page. Analysis name and owner, and the postscript file name, are text fields to be entered, all others are choices that are selected by clicking a button icon next to the desired value.

5.2.4 Remove TAP output page

This page provides a way to remove output files from TAP analysis jobs from a TAP user directory. Care must be taken when using this page to specify the correct Analysis name(s) and owner. Should this page be selected by mistake, one can always "back out" by clicking on the main menu link or using the Netscape "back" feature: Files will not be deleted until the button labeled "REMOVE cases" is clicked.

The inputs required by this page (Figure 5) are quite similar to those of the status check (Section 5.2.1). Note, however, that the default field for the analysis name is chosen such that it will not match any likely case names, to avoid accidental deletion of cases. If all cases are to be deleted, a blank input string has to be entered into the input field.

```
Remove TAP cases

   * Analysis Name:
 -  There are 2 Options:
 -- (1): remove all cases: Fill in a blank for the Analysis Name
 -- (2): remove one or more case(s): Analysis name(s) must be specified

   * Analysis owner:
```

Figure 5: Required inputs for the TAP remove menu page. Analysis name and owner are text fields to be entered.

Submitting this form will launch a UNIX job that will remove all output files from the specified analysis name(s) in the user's directory. Output from the unix command is echoed to the screen.

5.3 The csh user interface

The csh user interface is provided as a fallback option, should the graphical user interface be unavailable. It also provides the option for more experienced Unix and TAP users to automate some TAP functions.

The csh user interface is started with the `tap.csh` command (see Section 5.3.2). A basic help command (`tap.help`) may be issued before starting the csh user interface.

5.3.1 The tap.help command

This command is invoked by entering `tap.help`. It produces output to "standard output" (normally the screen), containing a list of all TAP commands, and current settings of TAP-related environment variables and command "aliases". A sample output is shown in Figures 6 and 7.

```
Theater Analysis Procedures (TAP) help:

Required initialization:
   need to "setenv TAPHOME ..." and "source $TAPHOME/setenv.csh"

Start TAP graphical user interface:
   tap.web

Start csh user interface:
   tap.csh - this changes your working directory
     Also needed when changing the TAP user name

csh user interface commands:
   tap.execute - execute a TAP analysis run
 (invoke without arguments for help)
   tap.map-plot - produce a plot from results of a TAP analysis run
 (invoke without arguments for help)
   tap.check - list all analysis names
 (optional arguments: produce a full list for those analysis names)
   tap.remove - Remove all output files for each analysis name
 (optional arguments: only remove those analysis names)
```

Figure 6: Sample output from the tap.help command – Part A: List of commands.

```
Your present TAP setup is:

your directory:/users/cwf/tap.demo/TAPHOME/users/tap

TAP-related environment variables:
PWD=/users/cwf/tap.demo/TAPHOME/users/tap
TAPgrib=/users/cwf/tap.demo/TAPHOME/DATA/GRIB
TAPfixed=/users/cwf/tap.demo/TAPHOME/splus/.Fixed
TAPaims=/users/cwf/tap.demo/TAPHOME/DATA/AIMS
TAPSTATS=/users/cwf/tap.demo/TAPHOME/error_stats/baseline
TAPHOME=/users/cwf/tap.demo/TAPHOME
TAPFORT=/users/cwf/tap.demo/TAPHOME/fortran
TAPC=/users/cwf/tap.demo/TAPHOME/C
OORTHOME=/users/cwf/tap.demo/TAPHOME/oort/fmtfilt

TAP-related aliases:
tap.check csh -f $TAPHOME/splus/check-cases.csh !* | \
 sed -e "s/<[/]*strong>//g" | sed -e "s/<[/]*h[1-9]>//g"
tap.csh source $TAPHOME/splus/tap.csh
tap.execute csh -f $TAPHOME/splus/tap.execute.csh
tap.help source $TAPHOME/splus/tap.help.csh
tap.map-plot csh -f $TAPHOME/splus/tap.map-plot.csh
tap.remove setenv prompt $prompt ; csh $TAPHOME/splus/remove-cases.csh !* ; unsetenv
tap.web netscape $TAPHOME/html/main.menu.html
```

Figure 7: Sample output from the tap.help command – Part B: TAP environment.

5.3.2 Starting the csh user interface

This command is invoked by issuing `tap.csh`. It will prompt for input from "standard input" (normally the keyboard), asking for the TAP user name. It will then change the current working directory to that user's TAP directory. Specifying an invalid TAP user name will result in repeated prompts for valid user names. This can be stopped by entering "quit" instead of a user name. A sample session is shown in Figure 8.

```
Starting the TAP csh user interface:
Enter your TAP user name (usually "tap"; "quit" if you want to quit now): tan
This is an invalid TAP user id: tan
Please try again
Enter your TAP user name (usually "tap"; "quit" if you want to quit now): tap
/users/cwf/tap.demo/TAPHOME/users/tap

Your current directory is now:/users/cwf/tap.demo/TAPHOME/users/tap
You need to repeat "tap.csh" to change TAP user names or
if you issue any "cd" commands during your csh session

Issue tap.help for basic TAP help
```

Figure 8: Sample session of the tap.csh command.

5.3.3 Status check of TAP analysis jobs

Command `tap.check` is provided for checking on the status of TAP analysis jobs in the user's directory. When invoked without command line arguments, it will list all analysis jobs in the user's directory, indicate whether they are still running or have completed, and list the contents of their timestamp files. (The command keys on the presence of the timestamp, which are files with the filenames extension `.timestamp`, and which contain information on the hostname, start and end date and times, and variables and levels analyzed, for an analysis job.) When invoked with optional command line argument(s), the examination is restricted to the analysis names specified on the command line (this corresponds to the Analysis Name input field in Figure 2). In this case `tap.check` provides a full listing of diagnostic information for each of the analysis names (it should thus be invoked in combination with the Unix `more` command or output redirection). See Section 5.2.1 for a discussion of the output from this command.

5.3.4 Execute TAP analysis jobs

The `tap.execute` command launches an analysis job. It differs from the corresponding graphical user interface command (Section 5.2.2) in two important respects: the analysis name is constructed internally and not specified by the user; more than one analysis job can

be spawned by specifying multiple options for any one of the four possible input parameters. To avoid simultaneous execution of multiple analysis jobs, the command script waits for the completion of each job before starting the next job (or exiting the script). Thus, it is best to run this command in the background (and redirect its output to a file for later examination). A typical sequence of commands is to invoke the command without any arguments to display its help message (Figure 9), and then to invoke it again with command line arguments, output redirection. and in the background:

```
tap.execute
tap.execute Mar95NE "( zuv rh )" large eta12 >&! test.out &
```

The command line arguments closely correspond to the choices of the TAP execute page (Section 5.2.2). In this example, two analysis jobs will be run: one is a height and wind analysis, the other a relative humidity analysis, both for the March 1995 case over the Northeast region, using the large analysis domain and the ETA 12-hour forecast as a background field. The analysis names are constructed from the selected options. In this example, they are `Mar95NE.zuv.large.eta12` and `Mar95NE.rh.large.eta12`.

```
tap.execute needs 2-4 args:
1: caseIds. one or more of  ( Mar95NE Mar95SE OpalNE OpalSE )
2: aschemes. one or more of  ( zuv z uv t rh ts )
3: grid.types. One or more of ( column small large ). Default: column
4: Background types. One or more of (climo eta12). Default: climo

NOTES:
(1) if giving more than one in any of the above, need to do it in this form
    (using the exact same quotes and spaces): "( column small )"
(2) run this script in the background (Put a "&" at the end of the command
    line) if you want to perform other tasks while tap.execute is running
```

Figure 9: Help message from the tap.execute command.

5.3.5 Plot a TAP plot

Tap plots are produced by the tap.map-plot commmand. Invoking the command without any command line arguments will display its help message (see Figure 10). The command line arguments closely correspond to the menu choices of the TAP plot page (Section 5.2.3).

5.3.6 Remove TAP output

Output files from TAP analysis jobs are removed by the `tap.remove` commmand. Invoking the command without any command line arguments will remove the output from all analysis jobs in the current directory. Optional command line arguments are names of cases to be

```
tap.map-plot needs 1-5 args:
1: Analysis name. You must know this name to plot it.
2: Variable. One of: "c(7,33,34)", 7, "c(33,34)", 11, 52
   corresponding to:   z + (u,v) , z,(u,v) , t , rh
   Default: "c(7,33,34)"
3: Level.  Pressure in mb, or "0" for surface plots
   Default: "500"
4: Type of plot.  One of "anv" (analysis values);
   "bgv" (background values); "inc" (Increments)
   Default: "anv"
5: Flag ("T" for yes, "F" for no) for overlaying observations
   Default: "T"
6: Postscript file name
   Default: "tap.ps"
```

Figure 10: Help message for the tap.map-plot command.

Generating a plot: less than 2 minutes
Generating an analysis:

Analysis type	large or small grid	column at domain center
Height and winds	30 minutes	6 minutes
Winds	20 minutes	5 minutes
Scalar (Temperature, RH)	14 minutes	5 minutes
Surface temperature	7-10 minutes	2 minutes

Table 1: TAP execution time estimates

removed. (This is analogous to the analysis name input field in Figure 5). When run in an interactive shell, this command will prompt the user for confirmation before removing any files.

5.4 TAP execution times

Table 1 lists wall clock times for the early prototype TAP. These estimates were obtained on a dedicated Sun workstation, using the test cases of the CWF test. This table is also accessible from the graphical user interface (the execution and plotting pages both provide links to this table).

5.5 Troubleshooting and error recovery

In the following, some possible error conditions and their likely causes and remedies are listed. This section is geared toward the graphical user interface; since the underlying problems will be the same in most cases for both user interfaces, it can also be used for the csh user

interface. In case these suggestions do not solve the problem, help should be sought from the TAP developers/software support.

Screen display shows the line "Can't cd to ..." on top: this is caused by an invalid Analysis Owner name. Correct the form and resubmit.

Ghostview produces an error message: "Warning: failed to allocate ... RGB cube." This occurs when other applications running on the workstation (such as Netscape) have already used up too many colors from the workstation color table. This may affect the colors displayed to the screen, but the message can otherwise be ignored.

Screen display shows TAP request is submitted, but nothing happens: there are several possible reasons.

1. The plotting job may still be executing. Plots should be displayed on the screen after a minute or two (see Table 1).
2. The analysis may still be executing. Use the Status check page to check for this. Execution times for analyses range from 2 - 30 minutes (see Table 1).
3. An invalid TAP user name is used. Make sure to check the top of the page shown after you submit the form for the "Can't cd to ..." error message, and correct the user name if needed.
4. The plotting job may have failed because an invalid analysis name was specified. Double-check your analysis name (use the Status check page if you forgot the name of your analysis).
5. The plotting job may have completed, but the output could not be displayed. If running Netscape on a different machine than the server, make sure to use the xhost command to add the server to your access list. If running the csh user interface, make sure the DISPLAY environment variable is set correctly. If all else fails, examine the user's directory for the presence of the postscript file, and examine the plotting job log file.
6. The plotting job may have failed because the analysis output is corrupted. See the discussion of analysis errors below.

Status check lists the Unix command, but produces no output: this is not an error, but reflects the fact that no analyses were found. Make sure you specified the correct Analysis owner and name.

Status check does not show the analysis: (even though the screen display shows the TAP Splus tap.execute request is submitted) There is a short delay between starting the Splus job and the creation of the timestamp file. If status check is used too soon after the submission of the request, the timestamp file may not be created yet. Another possibility is that the Analysis owner and name were incorrectly specified.

Status check gives a warning message about a job running in BATCH: This means that there apparently is an analysis job running in the user directory which was started

from the csh user interface. This is detected by the presence of a file (`BATCH.RUNNING`) in the user's directory. No analysis jobs should be started until this jobs has completed or has been aborted.

How do I kill an analysis job? This requires knowledge of the UNIX `ps`, `kill` commands. From the UNIX command prompt, the processes launched by the execute page (or the tap.execute command) must be located and killed. They can be identified by their process IDs, and/or their command names. Both the graphical user interface and the tap.execute script use the `Splus BATCH` command, which spawns an Splus process (this will appear as `Sqpe` in the listing produced by `ps`). As an alternative to kill, it is also possible to repeatedly remove (`tap.remove`) all output files from the running analysis job - this will lead to internal errors in the analysis job and its abnormal termination; however, note that the job may still be running in this case even though its `.timestamp` file no longer exists, in which case the Status check page will incorrectly indicate that there are no more jobs running.

How do I remove output files from old or killed analysis jobs? This can be done from the graphical (Section 5.2.4) and the csh user interface (`tap.remove`).

Causes of analysis job errors : Aside from internal execution errors, which should be referred to the TAP developers/software support, analysis jobs may terminate abnormally because of system-level problems:

1. Running out of memory. This will be indicated by an error message in the Splus job log file (produced by the Status check page): "Cannot allocate requested dynamic memory ...". Make sure no other memory intensive jobs are running on the system at the same time as the TAP analysis job. Make sure enough swap space is allocated to the system.
2. Running out of disk space. Use the `tap.remove` command to clean out user directories as needed.

6 Test Description

For purposes of the prototype tests, TAP is restricted to running a number of canned cases over prespecified regions and analysis domains. These have been chosen to present a typical sample of options of the envisioned operational TAP system. At present, the data sources are restricted to radiosondes for the upper air analyses, and surface station reports (SYNOP, Service A, and ship/buoy reports) for the surface analyses.

6.1 Analysis domains

A set of three nested analysis domains is used in two separate geographic regions, one over the Northeast and one over the Southeast United States. The outermost, large grid domain covers a region of approximately 1500 km on a side. Centered inside the outer region is the small domain, covering a region of approximately 750 km on a side. Both grids consist of

11 by 11 gridpoints. Finally, the column domain represents the grid column at the center grid point of the small analysis domain. Over the Northeast region, a polar stereographic projection basemap is used (with a reference longitude at 80° W), and the analysis domains are centered over southeast Pennsylvania (see Figure 11). Over the Southeast region, a Mercator projection basemap is used, and analysis domains are centered over Alabama (see Figure 12). The analysis grids were chosen to contain both land areas with dense data coverage and data sparse areas over water.

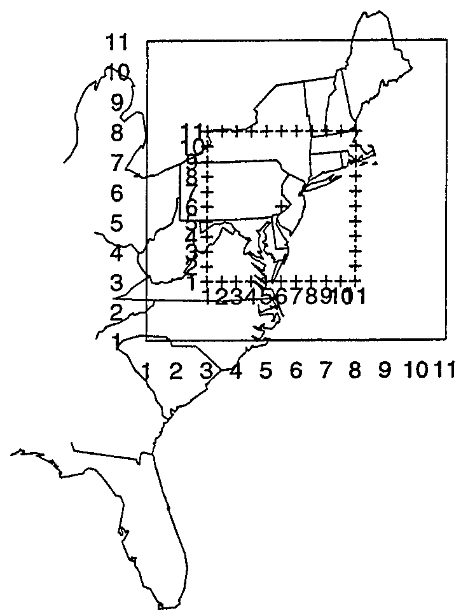


Figure 11: The three analysis domains over the Northeast region.

6.2 Test cases

6.2.1 March 1995 case

The analyses for this case are for 12 UTC 6 March 1995. The synoptic situation on that day was characterized by an upper-level shortwave passing through the Northeastern United States and Canada, embedded in a southwesterly upper-level flow. At the surface, this was accompanied by a weak low (central pressure around 1016 hPa) and rain over the Northeastern United States, snow over parts of Quebec and the Canadian Maritimes. This case represents a typical moderate to weak winter/early spring storm over the Northeastern United States. An example TAP 850 mb height and wind analysis is shown in Figure 13. Over the southeast region, the situation is dominated by a ridge (see Figure 14).

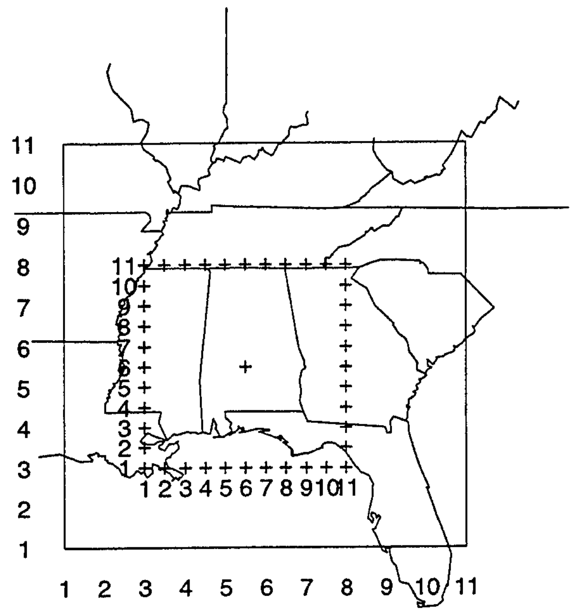
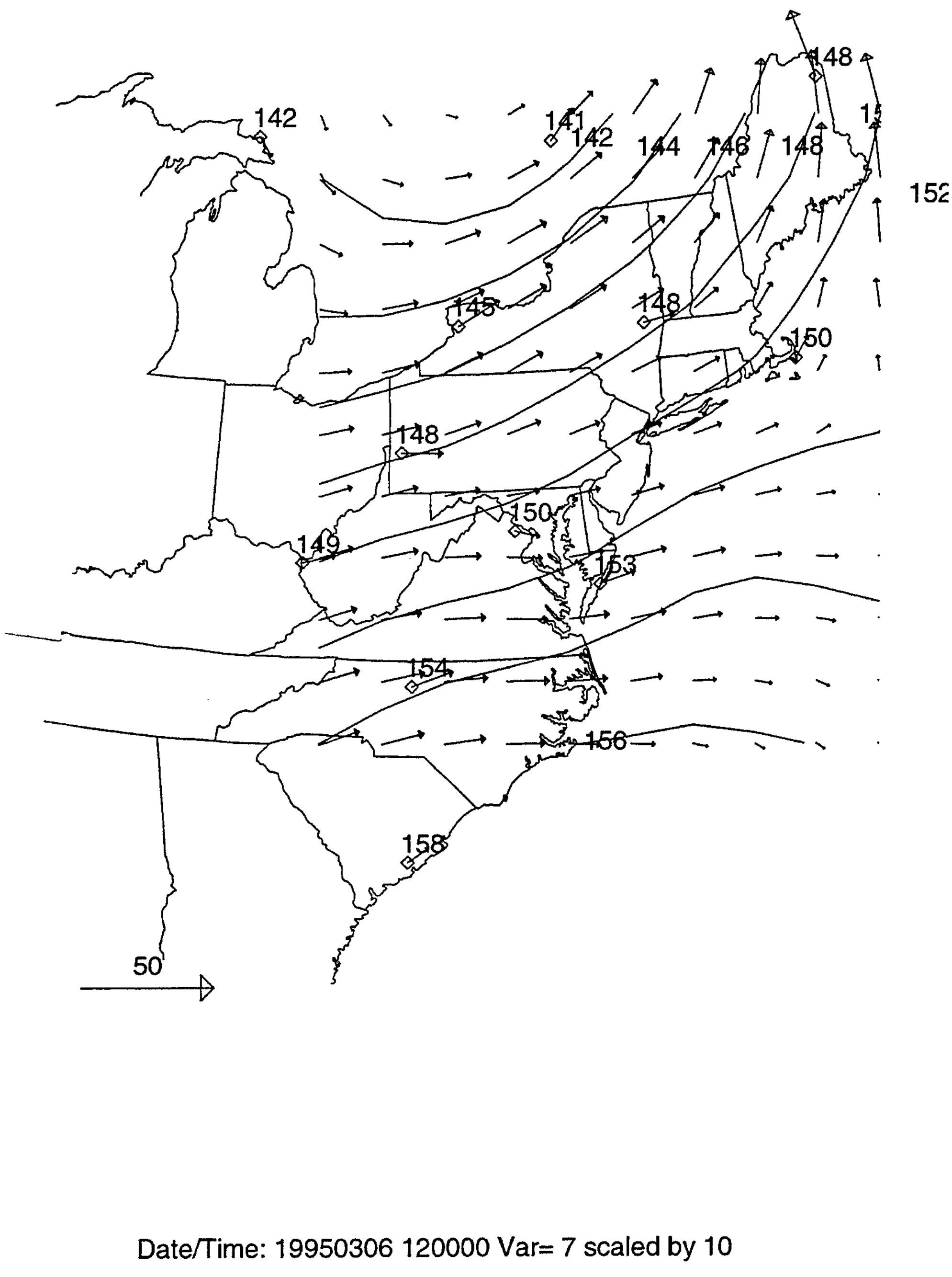


Figure 12: The three analysis domains over the Southeast region.

6.2.2 Opal (October 1995) case

The analyses for this case are for 00 UTC 5 October 1995, approximately 1 hour after Hurricane Opal made landfall in the Florida panhandle (Figure 15). Opal was a strong hurricane, causing over $3 billions of damage and over 20 deaths. It thus represents an extreme weather event. Unfortunately, because of the extreme weather conditions, key radiosonde reports are missing for this case at some or all of the levels. The results of the analysis thus depend strongly on whether the ETA 12-hour forecast or climatology is being used as a background field. In particular, if the column analysis domain is chosen over the Southeast region, there will be no radiosonde reports in the data window surrounding the analysis point at some of the level. The analysis value in this case is simply equal to the background value. Over the Northeast region, there are weak shortwave features embedded in a generally southwesterly flow (Figure 16).

Mar95NE.zuv.large.eta12 anv Level= 850 Var= 7 33 34



Date/Time: 19950306 120000 Var= 7 scaled by 10

Figure 13: TAP analysis for height and winds at 850 mb over the Northeast region for the March 1995 case.

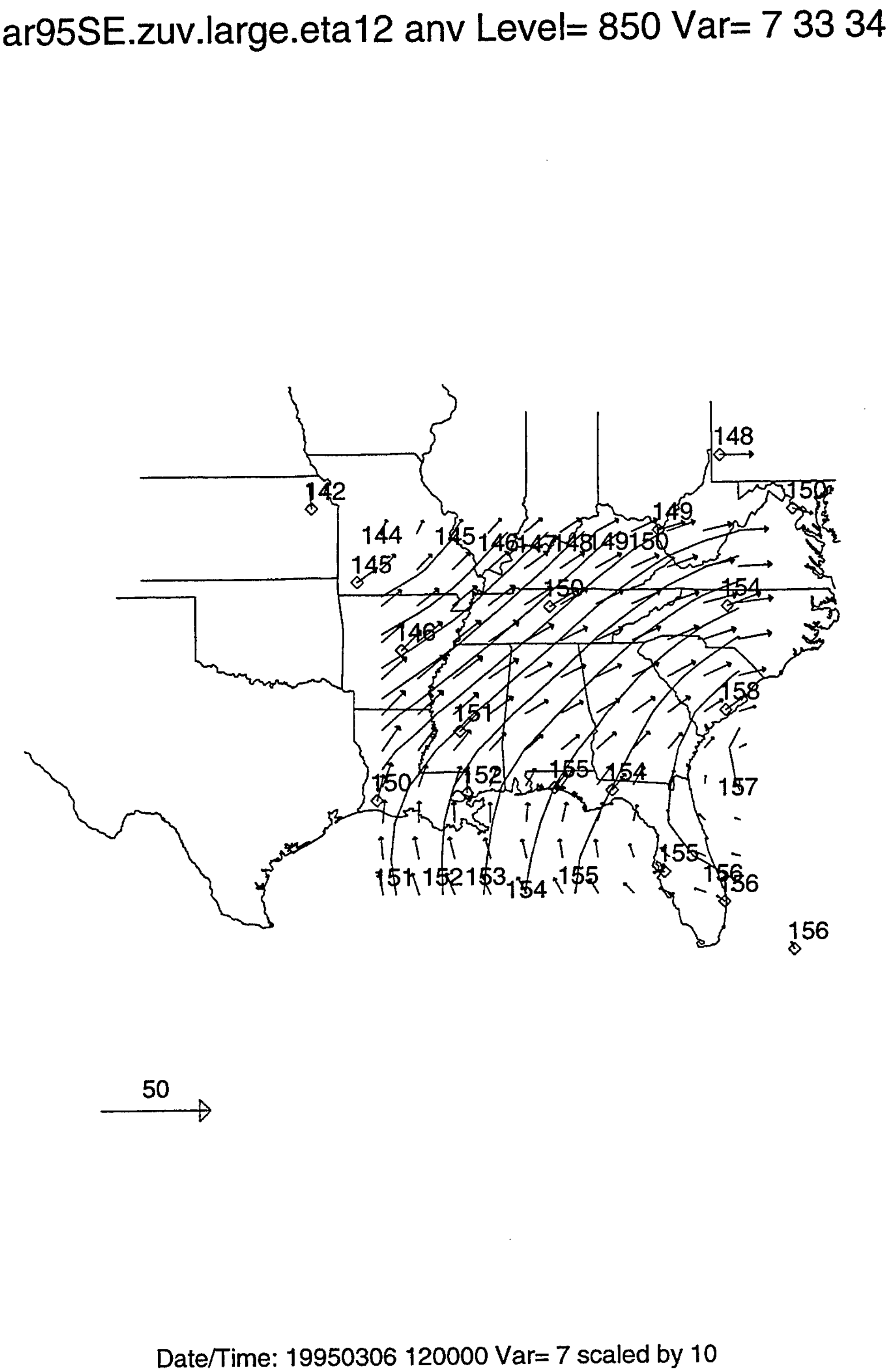
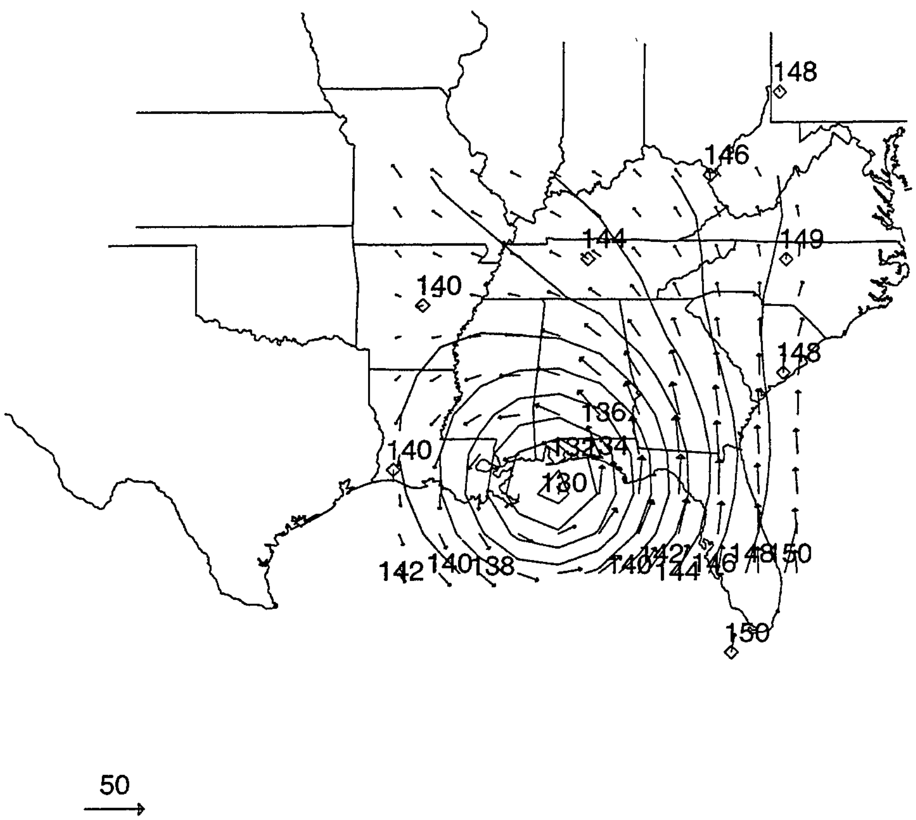


Figure 14: TAP analysis for height and winds at 850 mb over the Southeast region for the March 1995 case.

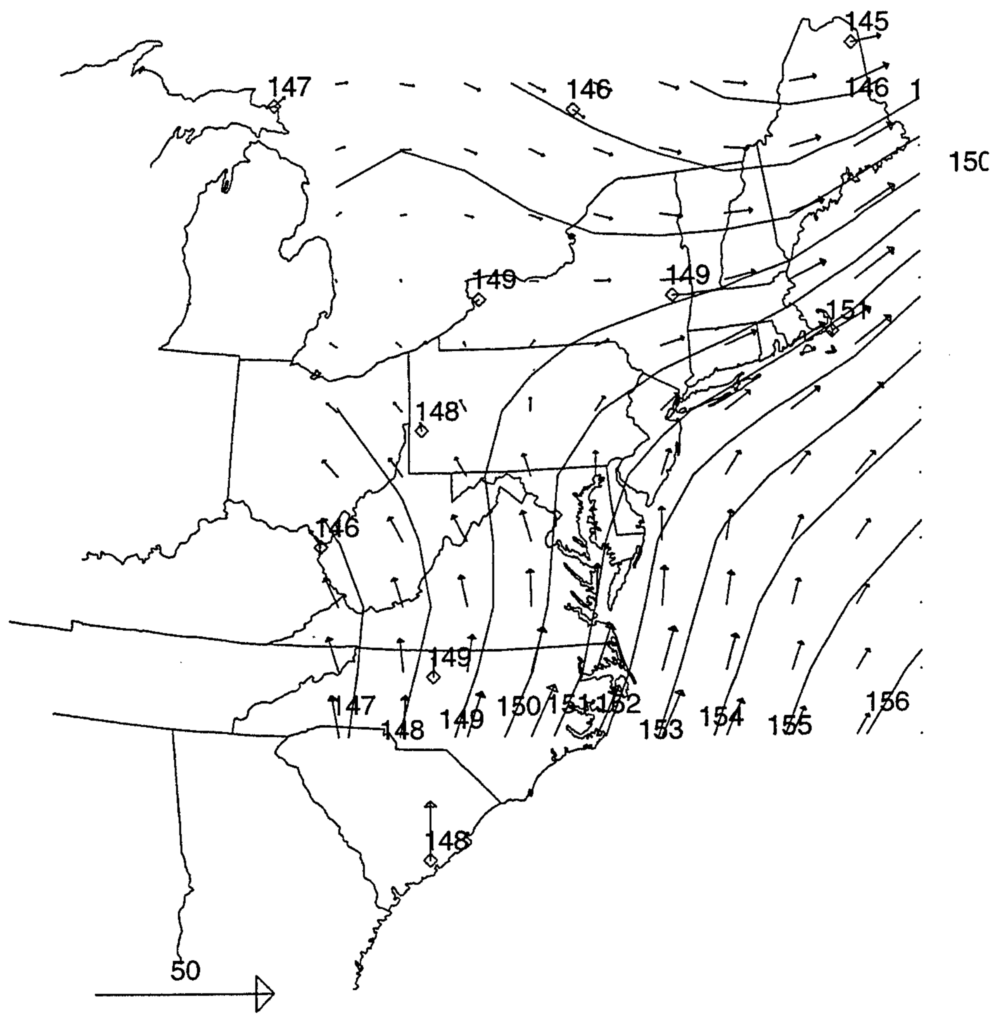
OpalSE.zuv.large.eta12 anv Level= 850 Var= 7 33 34



Date/Time: 19951005 0 Var= 7 scaled by 10

Figure 15: TAP analysis for height and winds at 850 mb over the Southeast region for the Opal case.

OpalNE.zuv.large.eta12 anv Level= 850 Var= 7 33 34



Date/Time: 19951005 0 Var= 7 scaled by 10

Figure 16: TAP analysis for height and winds at 850 mb over the Northeast region for the Opal case.

C User's Manual for Evaluation at the Air Force Weather Agency

User's Manual prepared for the quasi-operational evaluation of TAP at the Air Force Weather Agency (AFWA) in Omaha, Nebraska.

Contents

List of Figures

1 Introduction

The Theater Analysis Procedures (TAP) are being developed to provide a meteorological analysis capability that will reside on a computer workstation. The prototype system is being evaluated by personnel at the Air Force Weather Agency (AFWA, formerly known as Global Weather Central - GWC). The evaluation is designed to assess potential TAP applications, in particular for intialization of the MM5 mesoscale model.

2 Scope

2.1 Identification

This document is the User's Guide for the prototype evaluation of the TAP at the AFWA. It applies to the version of TAP after the basic TAP development effort. Refer to section 2.3 for a reader's guide.

2.2 System overview

2.2.1 Objectives of the TAP Project

The TAP project primary objective is to develop robust analysis procedures to support the tactical user. These analysis procedures provide stable meteorological products for end users.

TAP is modular, and capable of utilizing a variety of background and data sources. This capability allows TAP to adapt to different theater meteorological support systems (TMSSs), run on different platforms and satisfy different user requirements. TAP is configurable to a range of requirements, from first-in stand-alone capability to full Theater Weather Central (TWC) support.

2.2.2 Major Functions of TAP

The function of TAP is to use the optimal interpolation technique to combine background (i.e. *a priori*) information with observations of diverse type, quality, and density to produce analyses of meteorological fields. The TAP analysis configurations are optimized to initialize NWP models and to provide input for tactical decision aids (TDAs). The implementation at GWC is tailored to the initialization of the MM5 mesoscale model from large-scale model output after preprocessing by the MM5 `DATAGRID` program.

2.2.3 Management and Technical Constraints

TAP is state of the art, but not experimental. TAP is designed to work every time. In measuring the quality of TAP, robustness of the system is weighted heavily.

2.3 Document overview

This document gives a brief overall description of the TAP project, and provides instructions for installing and running the TAP system.

At a minimum, personnel should read Sections 5.1 and 5.2 for basic operation of the system. Installation is covered in Section 4, customization in Section 6, and troubleshooting in Section 5.3.

3 Referenced documents

For more detailed TAP documentation, see also the System and Interface requirements specifications and design documents (document `srs`, document `sdd`, document `irs`, and document `idd`). The installation procedure is additionally described in file `install.procedure` [1] in directory `$TAPHOME/install`.

4 Installing TAP

4.1 Required system software

The following system software is assumed to be preinstalled on the host system:

- Operating system
- Fortran 90 compiler
- C compiler

4.2 Installing supporting software

4.2.1 Splus

The TAP system development effort made extensive use of the Splus software package, and some functions of the present version of the software are still implemented in Splus. Splus is a proprietary software package [2] for interactive data analysis and display that is provided with its own storage medium and installation instructions. The location of the Splus software (both the location of the Splus Home directory and the directory of the executable Splus script) must be recorded for later use in the TAP installation.

4.2.2 Netscape server and browser

An earlier prototype of the TAP software made use of a graphical user interface based on the HTML protocol, requiring the installation and configuration of an HTML server and browser (such as Netscape). This is no longer required for the version used at AFWA, since there will be limited need for user interaction.

[1] All filenames included in the TAP installation tape are identified by the following font: `filename`.

[2] More information available on the Internet at `http://www.mathsoft.com/splus/`.

4.2.3 Others

Other software packages have been assembled onto a tar tape and can be directly restored from tape to disk. Necessary for operation of the basic TAP system are: LAPACK and BLAS source code files (only those routines actually used by TAP are included here), which are automatically loaded to disk by untarring the TAPHOME directory.

Additional software are used by ancillary functions of TAP, but are not included on the tar tape to simplify installation: GNU Make and related utilities to facilitate compilation; Ghostview, xv, and related utilities for display of graphical output; to support the generation of postscript documents from the LaTeX source files, the LaTeX package is also needed; similarly, generation of HTML versions of the documents requires the Latex2html package.

Additional utilities and programs, which are only needed for the software development environment, are also not included in the basic package: RCS for version control, Gnu Emacs for editing of files and system-level interactive use of the TAP software.

4.3 Installing the TAP software

The TAP software and ancillary datasets are all stored together under the `TAPHOME` directory. The entire directory tree structure is stored on a tar tape, and installation only requires transfer from the tar tape onto the appropriate directory on the disk of the host system. For example, if the desired disk location is `/users/afwa/tap.demo/`, then the required Unix commands are

```
cd /users/afwa/tap.demo
tar -xvf /dev/rmt1 TAPHOME |& tee tar.out
```

where `/dev/rmt1` denotes the tape drive on which the tar tape is mounted. Upon completion of this command, a new directory `TAPHOME` is created, which contains all the TAP software and ancillary data. As described in Section 5.1, the full path name of this directory needs to be stored in the environment variable `$TAPHOME`.

After transfer of the data to disk, some editing of files is required to change path names of certain system and supporting software. A cshell script (`install.csh`) in the `$TAPHOME/install` directory is provided for this purpose. The script creates and executes a file of editor commands (for the UNIX `sed` editor), based on user responses to queries for locations of the various software packages and system binaries. Default values are provided for most of these. Refer to file `install.procedure` in directory `$TAPHOME/install` for up-to-date instructions.

4.4 Adding or removing TAP users

TAP is set up to support multiple users. Separate disk areas are provided for each user for storage of output and intermediate files. Adding a user, (for example: `bob`) simply requires the addition of subdirectories by the commands:

```
mkdir $TAPHOME/users/bob ; mkdir $TAPHOME/users/bob/.Data
```

Conversely, removing this user is accomplished by

```
"rm" -fr $TAPHOME/users/bob
```

For the usual, single-user mode of operation, the user name `tap` is preinstalled on the system.

5 Running TAP

A shell script interface for running the TAP system is provided for the evaluation at AFWA, which consists of shell scripts which are invoked by the user from the system prompt. The HTML-based graphical user interface developed for beta-testing of the early prototype is not included in the AFWA installation, since only limited user interaction is anticipated.

5.1 Required initializations for the Unix session

Before invoking TAP through the shell user interface, the environment variable `$TAPHOME` must be set to the location of the tap software, e.g. in the above example it would be

```
setenv TAPHOME /users/afwa/tap.demo/TAPHOME
```

Further environment variables and command aliases are then set by issuing the command

```
source $TAPHOME/setenv.csh
```

Note: Both of these commands could be included in the user's `.cshrc` file, so that they will be executed automatically upon login, or included in a batch job script for batch execution of TAP.

A simple help command (`tap.help`) is provided that lists all available tap commands and the user's TAP configuration. It is described in more detail below.

5.2 The shell user interface

The shell user interface is started with the `tap.csh` command (see Section 5.2.2). A basic help command (`tap.help`) may be issued before starting the csh user interface.

5.2.1 The tap.help command

This command is invoked by entering `tap.help`. It produces output to "standard output" (normally the screen), containing a list of all TAP commands, and current settings of TAP-related environment variables and command "aliases". A sample output is shown in Figures 1 and 2.

```
Theater Analysis Procedures (TAP) help:

Required initialization:
   need to "setenv TAPHOME ..." and "source $TAPHOME/setenv.csh"

Start TAP graphical user interface:
   tap.web

Start csh user interface:
   tap.csh - this changes your working directory
     Also needed when changing the TAP user name

csh user interface commands:
   tap.execute - execute a TAP analysis run
 (invoke without arguments for help)
   tap.map-plot - produce a plot from results of a TAP analysis run
 (invoke without arguments for help)
   tap.check - list all analysis names
 (optional arguments: produce a full list for those analysis names)
   tap.remove - Remove all output files for each analysis name
 (optional arguments: only remove those analysis names)
```

Figure 1: Sample output from the tap.help command – Part A: List of commands.

```
Your present TAP setup is:

your directory:/users/afwa/tap.demo/TAPHOME/users/tap

TAP-related environment variables:
OORTHOME=/users/afwa/tap.demo/TAPHOME/oort/fmtfilt
PWD=/users/afwa/tap.demo/TAPHOME/users/tap
TAPaims=/users/afwa/tap.demo/TAPHOME/DATA/AIMS
TAPBLAS=/users/afwa/tap.demo/TAPHOME/BLAS
TAPC=/users/afwa/tap.demo/TAPHOME/C
TAPfixed=/users/afwa/tap.demo/TAPHOME/splus/.Fixed
TAPFORT=/users/afwa/tap.demo/TAPHOME/fortran
TAPgrib=/users/afwa/tap.demo/TAPHOME/DATA/GRIB
TAPHOME=/users/afwa/tap.demo/TAPHOME
TAPLAPACK=/users/afwa/tap.demo/TAPHOME/LAPACK
TAPSTATS=/users/afwa/tap.demo/TAPHOME/error_stats/baseline
TAPwork=/users/afwa/tap.demo/TAPHOME/users/tap

TAP-related aliases:
tap.check csh -f $TAPHOME/splus/check-cases.csh !* | \
 sed -e "s/<[/]*strong>//g" | sed -e "s/<[/]*h[1-9]>//g"
tap.csh source $TAPHOME/splus/tap.csh
tap.dev.csh source $TAPHOME/splus/tap.dev.csh
tap.execute csh -f $TAPHOME/splus/tap.execute.csh
tap.help source $TAPHOME/splus/tap.help.csh
tap.map-plot csh -f $TAPHOME/splus/tap.map-plot.csh
tap.remove setenv prompt $prompt ; csh $TAPHOME/splus/remove-cases.csh !* ; \
                        unsetenv prompt
tap.web netscape $TAPHOME/html/main.menu.html
```

Figure 2: Sample output from the tap.help command – Part B: TAP environment.

5.2.2 Starting the shell user interface

This command is invoked by issuing `tap.csh`. It will prompt for input from "standard input" (normally the keyboard), asking for the TAP user name. It will then change the current working directory to that user's TAP directory. Specifying an invalid TAP user name will result in repeated prompts for valid user names. This can be stopped by entering "quit" instead of a user name. A sample session is shown in Figure 3.

```
Starting the TAP csh user interface:
Enter your TAP user name (usually "tap"; "quit" if you want to quit now): tan
This is an invalid TAP user id: tan
Please try again
Enter your TAP user name (usually "tap"; "quit" if you want to quit now): tap
/users/afwa/tap.demo/TAPHOME/users/tap

Your current directory is now:/users/afwa/tap.demo/TAPHOME/users/tap
You need to repeat "tap.csh" to change TAP user names or
if you issue any "cd" commands during your csh session

Issue tap.help for basic TAP help
```

Figure 3: Sample session of the tap.csh command.

For batch execution of TAP, the commands executed by the `$TAPHOME/splus/tap.csh` script could be included in the batch job instead, eliminating the need for user interaction.

5.2.3 Status check of TAP analysis jobs

Command `tap.check` is provided for checking on the status of TAP analysis jobs in the user's directory. When invoked without command line arguments, it will list all analysis jobs in the user's directory, indicating whether they are still running or have completed, and list the contents of their timestamp files. (The command keys on the presence of files with the filename extension `.timestamp`, which contain information on the hostname, start and end date and times, and variables and levels analyzed, for an analysis job.)

When invoked with optional command line argument(s), the examination is restricted to the analysis names specified on the command line. In this case `tap.check` provides a full listing of diagnostic information for each of the analysis names (it should thus be invoked in combination with the Unix `more` command or output redirection): the size of all data files for that analysis, and the contents of the log file. This display is useful for debugging purposes, but should not usually be required.

5.2.4 Execute TAP analysis jobs

The `tap.execute` command launches an analysis job. To avoid simultaneous execution of multiple analysis jobs, the command script waits for the completion of each job before

starting the next job (or exiting the script). Thus, it is best to run this command in the background (and redirect its output to a file for later examination). A typical sequence of commands is to invoke the command without any arguments to display its help message (Figure 4), and then to invoke it again with command line arguments, output redirection, and in the background:

```
tap.execute
tap.execute -l mm5 mm5.2 mm5.3 mm5.4 mm5.5 datagrid.out \
            correl.ctl.file nes.ctl.file RunA >&! RunA.job.out &
```

In the example provided above, the second line starts a TAP analysis job using a DATAGRID output file with the default set of customization files and analysis name "RunA".

The -s command line switch causes the Splus version of the analysis code to be used instead of the Fortran 90 version. At present, most functions have been translated to C or Fortran 90, but there are some remaining functions that are only implemented in Splus: the initialization of the error statistics and various control parameters, the preprocessor (in which error standard deviations and background values are added to analysis grid points and observations as needed), and the Splus postprocessor (in which analysis values are output in a gridded format more suitable for graphical display). For the AFWA implementation, the ingest and partial preprocessing of the background and observations, and the postprocessing of the analysis values into the DATAGRID format, are all accomplished by Fortran 90 modules. The -s command line switch does not affect any of these processing steps, but only the analysis part of the code itself, in which preprocessed background and observation values are combined to provide analysis values. The Splus version of this part of the processing should only be required if the OI quality control is to be used, since this part of the algorithm is not yet implemented in Fortran.

The -l command line switch is used in cases when local versions (*i.e.*, residing in the user's directory) of the Splus and/or Fortran executables, or the main Splus execution script, should be used instead of the default versions of the TAP installation. This is useful for testing customizations of the TAP algorithm before implementing any changes in the TAP installation (see Section 6).

The first five arguments are all provided for compatibility with non-AFWA implementations of TAP, for which the analysis date/time, grid, variables and levels are all specified independently of the input background field. The casedate is used to specify the date and time, the grid.region selects one of a number of predefined grids, and the grid.type one of three available analysis grid options ("large": the entire predefined grid; "small": a nested grid at the approximate center of the large grid with half the grid spacing; or "column": the single center grid point). The cvars.name selects one of a predefined set of combinations of variable codes and values for the mass/wind flag. Similarly, the clevs.name selects one of a predefined set of levels and choices of vertical coordinates. The default sets of choices are contained in files d.agrids.table, d.avars.table, and d.alevs.table (all in directory $TAPHOME/splus).

For the AFWA implementation, all these parameters are determined from the input background field, *i.e.* the DATAGRID output file. For this setup, the casedate string must be specified as "mm5", and the remaining four arguments may be set to any string (they will be ignored).

```
tap.execute needs 2-9 args:
-s optional switch: use Splus analysis
-l optional switch: use local local.Sqpe, tap_anal, xecute.input if available
1: casedate(s) in the form of yyyymmddhh, or "mm5" for mm5 runs
2: grid.region. one or more of  ( NE SE [others] ) - ignored for mm5 runs
3: grid.types.  One or more of ( column small large [others]). Default: column
                - ignored for mm5 runs
4: cvars.names. One or more of ( zuv z uv t rh [others] ). Default: all
                Default for mm5 runs: "mm5" - same as in DATAGRID input file
5: clevs.names. One or more of ( s ua u1000 [others] ). Default: ua
                Default for mm5 runs: "mm5" - same as in DATAGRID input file
6: Background types. One or more of ( climo eta12 ). Default: climo
                for mm5 run: Should be DATAGRID input file name
7: correl.ctl.file. Default: correl.ctl.table
8: nes.ctl.file. Default: nes.ctl.table
9: Beginning of caseName (cases named "name", "name2", etc).
   Default: generated from grid.region, type, etc

NOTES:
(1) if giving more than one in any of the above, need to do it in this form
    (using the exact same quotes and spaces): "( column small )"
(2) run this script in the background (Put a "&" at the end of the command
    line) if you want to perform other tasks while tap.execute is running
(3) Without -l, a symbolic link to $TAPHOME/splus/local.Sqpe is created
    before execution and removed thereafter.  Any local copies
    will be destroyed
(4) Without -l, symbolic links to $TAPHOME/fortran/ are created
    before execution and removed thereafter for all f90 executables
    (tap_anal, tapprep, tap_obsing, tappostp).  Any local copies
    will be destroyed
```

Figure 4: Help message from the tap.execute command.

The background type is used to construct the input filename of the background field. For the AFWA implementation, the full pathname is constructed from the concatenation of the the `$TAPgrib` environment variable, a "/", and the background type string.

The `correl.ctl.file` is a file containing Splus statements which are executed before the ingest of the background; the `nes.ctl.file` is a file with statements to be executed after the background ingest, and before the observation ingest. Both files can be used to customize aspects of the TAP execution (see Section 6).

The name of the analysis must be a legal UNIX file name and a legal Splus variable name. This is always guaranteed if it begins with a letter and contains only alphanumeric characters. It is important to make a note of the name of the analysis, because this name is needed to generate plots.

Warning: No other analysis jobs should be running in the TAP user directory when this command is invoked. The `tap.check` command can be used to check for this.

Warning: Reusing a name will overwrite a pre-existing analysis.

5.2.5 Plot a TAP plot

A number of plotting functions have been written to support the development and testing of TAP, and a simple command interface is provided for producing basic plots of analysis and overlaid observation values. Plots will be plots of one or more variables at the analysis grid points, at one of the analysis levels, over a map background of the analysis domain. Values of observations used in the analysis may optionally be overlayed over the plot. These plots are produced by the tap.map-plot commmand. Invoking the command without any command line arguments will display its help message (see Figure 5).

```
tap.map-plot needs 1-5 args:
1: Analysis name. You must know this name to plot it.
2: Variable. One of: "c(7,33,34)", 7, "c(33,34)", 11, 52
   corresponding to:   z + (u,v)  , z,(u,v)  , t , rh
   Default: "c(7,33,34)"
3: Level.  Pressure in mb, or "0" for surface plots
   Default: "500"
4: Type of plot.  One of "anv" (analysis values);
   "bgv" (background values); "inc" (Increments)
   Default: "anv"
5: Flag ("T" for yes, "F" for no) for overlaying observations
   Default: "T"
6: Postscript file name
   Default: "tap.ps"
```

Figure 5: Help message for the tap.map-plot command.

Upon completion of the plot job, a postscript viewer (ghostview) is invoked to display the graphical output to the screen. The ghostview window contains pull down menus for printing

a hard-copy of the output, saving it to a file with a different name, magnification/reduction, and numerous other display options.

The name of the analysis must be specified. It must be the name of a previously completed analysis job.

For variables other than wind vectors, contour plots are drawn if the analysis name is an analysis over the small or large domain. If the analysis is for a single column, the value is printed at the analysis point location. The code numbers displayed with the choice of variables refer to the code numbers assigned to each variable within TAP. These code numbers are also displayed on the title of the plots. Selecting either a variable or a level that is not present in an analysis will produce a message (in the postscript output file and the display generated from it) listing the available levels and/or variables.

The plot type options identify what type of values are displayed: either analyzed values, background values (interpolated to the analysis locations from either the forecast or climatology first guess), or increments (the difference between the first guess and the analysis values). For increment plots, observation plots will be those of observation increments (i..e., the difference between the observed value, and the background value interpolated to the observation location).

Note:

1. The present version of the software uses a map database appropriate for the United States. For plots over other regions of the world, map databases are needed which are not part of the standard distribution of the Splus software. As a workaround for this problem, the `map-plot.script` can be modified to specify "database=NULL" as one of the arguments to "plot.std.map", which will suppress drawing of any map background.

2. For the AFWA implementation, the Splus objects needed for plotting are only generated if the "debug" flag is set to "T" (see Section 6).

5.2.6 Remove TAP output

Output files from TAP analysis jobs are removed by the `tap.remove` commmand. Invoking the command without any command line arguments will remove the output from all analysis jobs in the current directory. Optional command line arguments are names of cases to be removed. When run in an interactive shell, this command will prompt the user for confirmation before removing any files.

5.3 Troubleshooting and error recovery

In the following, some possible error conditions and their likely causes and remedies are listed.

Ghostview produces an error message: "Warning: failed to allocate ... RGB cube." This occurs when other applications running on the workstation (such as Netscape) have already used up too many colors from the workstation color table. This may affect the colors displayed to the screen, but the message can otherwise be ignored.

TAP job started, but nothing else happens: there are several possible reasons.

1. The plotting job may still be executing. Plots should be displayed on the screen after a minute or two.
2. The analysis may still be executing. Use the `tap.check` command to check for this. Execution times for analyses will depend on the size of the grid, the number of observations, and the host machine.
3. The plotting job may have failed because an invalid analysis name was specified. Double-check your analysis name (use the `tap.check` command if you forgot the name of your analysis).
4. The plotting job may have failed because the "debug" flag was not set (see Section 6). In this case the Splus postprocessor, which creates the Splus data objects needed for plotting, is not run.
5. The plotting job may have failed because the analysis output is corrupted. See the discussion of analysis errors below.
6. The plotting job may have completed, but the output could not be displayed. This could be because the DISPLAY environment variable is set incorrectly, or ghostview is not installed on the system. If all else fails, examine the user's directory for the presence of the postscript file, and examine the plotting job log file.

Status check produces no output: this is not an error, but reflects the fact that no analyses were found. Make sure you specified the correct analysis name.

Status check does not show the analysis: (even though the TAP analysis job has started) There is a short delay between starting the Splus job and the creation of the timestamp file. If status check is used too soon after the submission of the request, the timestamp file may not be created yet. Another possibility is that the analysis name was incorrectly specified.

Warning message about a job running in BATCH: This means that there apparently is an analysis job running in the user directory which was started from the shell user interface. This is detected by the presence of a file (`BATCH.RUNNING`) in the user's directory. No analysis jobs should be started until this job has completed or has been aborted, and the `BATCH.RUNNING` file has been removed. (The `tap.execute` command will not start a TAP analysis job if the file `BATCH.RUNNING` exists in the user's directory.)

tap.check incorrectly indicates jobs are still running: Since the `tap.check` command does not actually examine the jobs running on the system, but only examines the presence and contents of the `.timestamp` and `BATCH.RUNNING` files, it can produce incorrect results under certain conditions. The most likely reason is that a previous analysis job terminated abnormally, in which case the file `BATCH.RUNNING` may not have been removed as is normally the case, and/or the `.timestamp` file does not contain a line indicating the ending date and time of the analysis job.

How do I kill an analysis job? This requires knowledge of the UNIX `ps, kill` commands. From the UNIX command prompt, the processes launched by the `tap.execute` command must be located and killed. They can be identified by their process IDs, and/or their command names. The `tap.execute` script uses the `Splus BATCH` command, which spawns an Splus process (this will appear as `Sqpe` in the listing produced by `ps`). As an alternative to kill, it is also possible to repeatedly remove (`tap.remove`) all output files from the running analysis job - this will eventually lead to internal errors in the analysis job and its abnormal termination; however, note that the job may still be running in this case even though its `.timestamp` file no longer exists, in which case the `tap.check` command will incorrectly indicate that there are no more jobs running.

How do I remove output files from old or killed analysis jobs? This can be done using the `tap.remove` command.

Causes of analysis job errors : Aside from internal execution errors, analysis jobs may terminate abnormally because of system-level problems:

1. Running out of memory. This will be indicated by an error message in the Splus job log file (produced by the Status check page): "Cannot allocate requested dynamic memory ...". Make sure no other memory intensive jobs are running on the system at the same time as the TAP analysis job. Make sure enough swap space is allocated to the system.
2. Running out of disk space. Use the `tap.remove` command to clean out user directories as needed.

6 Customizing TAP

The instructions for running TAP so far only apply to the baseline settings of analysis domain, variables, and levels, of adjustable parameters, and the baseline set of algorithms. Instructions are provided here for relaxing all of these restrictions.

6.1 Adaptation to the AFWA environment

For the AFWA implementation, TAP is inserted in the MM5 preprocessing procedure. In the usual MM5 preprocessing procedure, the model grid domain and certain fixed fields are prepared by running program TERRAIN; in the next step, a large-scale gridded field (such as a global analysis or forecast) available on pressure surfaces is interpolated horizontally to the model gridpoints by running program DATAGRID; a successive corrections analysis program `RAWINS` is then used to modify the DATAGRID output, based on radiosonde observations; finally, the pressure level fields are interpolated vertically to the MM5 model levels by program INTERP. Some additional variable transformations and initialization steps are performed by INTERP, as well.

In the present version of the AFWA implementation, TAP is inserted in the MM5 preprocessing in place of the `RAWINS` program. The input to TAP consists of the file created

by DATAGRID, which provides the information needed to determine the TAP analysis grid, levels, and variables, and the values of the first guess or background field; and observation data files. In the present version of the software, radiosonde observations are read in from ASCII files. There are some system level settings that need to be configured for the proper execution of TAP:

Input background field: For the AFWA implementation, the full pathname of the input file containing the `DATAGRID` output is constructed from the concatenation of the `$TAPgrib` environment variable, a "/", and the background type string provided as a command line argument to `tap.execute`. If the directory will not change from one run to the next, it could be included in the `setenv.csh` file; otherwise, an appropriate `setenv TAPgrib` command must be issued before `tap.execute` is invoked. The DATAGRID program must be run for the "extended domain", so that observation increments can be computed without extraoplation of the background field data. The TAP analysis is performed on the unexpanded, "coarse-grid" domain (if the MM5 is run with nested grids, the coarse-grid values are interpolated to the finer mesh grids in the MM5 INTERP program).

Input observation filename: In the current version of the code, the filename of the input Raob data is is constructed from the concatenation of the `$TAPaims` environment variable, the string "/raob.", and the analysis date/time obtained from the `DATAGRID` output file (in the form "yymmddhh"). If the directory will not change from one run to the next, it could be included in the `setenv.csh` file; otherwise, an appropriate `setenv TAPaims` command must be issued before `tap.execute` is invoked.

Work space directory: The environment variable `$TAPwork` is used to specify a directory in which to store work files created by TAP. If it is not set, it defaults to the directory from which tap.execute is invoked (usually `$TAPHOME/users/tap`).

Analysis output filename: In the current version of the code, the filename of the TAP output (in the `DATAGRID` format) is constructed from the concatenation of the analysis name, and the string "tap.dgoutput". The file is placed in the work space directory (see above). This file should be used as input to the INTERP program. Note that it is identical in all respects to the DATAGRID file used as input to TAP, except that TAP analysis values have been placed in the coarse domain grid points of the geopotential height, winds, temperature, and relative humidity fields.

For testing the effect of changes in input background or observation data, the default settings of the environment variables described above must be changed to use the desired input files instead.

6.2 Customizing tunable parameters

The baseline set of adjustable parameters are initialized at the beginning of the TAP execution script. The TAP execute script reads two files that allows these initial settings to be overridden at execution time: a `correl.ctl.file` and a `nes.ctl.file` customization file.

The first of these is read in before the background ingest, and can be used to override default settings of parameters defined at that stage of the processing. The second is read in after the background ingest, but before the observation ingest, and can be used to override the settings of the remaining parameters. Default versions of each of these files (`correl.ctl.table` and `nes.ctl.table`) are stored in the `$TAPHOME/splus` directory. Different filenames can be specified as command line arguments to `tap.execute`. If the specified files exist in the working directory, *i.e.* the TAP user's directory, they are used; otherwise, the files are read in from the `$TAPHOME/splus` directory.

Customization can thus be performed either for the entire TAP installation (by modifying or creating new versions of control files in the `$TAPHOME/splus` directory), or for specific users (by creating new versions of control files in the TAP user directories).

The tunable parameters are stored in Splus data objects (so-called "list" structures), one each for controlling the computation of the error correlations (correl.ctl), the quality control procedures (qc.ctl), and the normal equation solver (nes.ctl). Some additional Splus data objects are used to control aspects of the data ingest and selection. A complete description of these data structures is provided as part of the software documentation. A brief summary of the most important parameters is given in the following. Examples of Splus statements to be included in the customization files for modifying their values are given in the default versions of the files.

6.2.1 correl.ctl: parameters controlling the error correlation computation

In the following, the names of the parameters are given, along with their default value and a brief explanation:

distmax (= -1): The maximum distance (expressed as the cosine of the angular great circle distance) beyond which all correlations are considered zero.

dothin (= T): Logical flag controlling whether to thin vertical profiles from radiosondes in the preprocessor. Used to prevent ill-conditioning of the matrix.

dpmax (= 30000): Maximum vertical separation (in Pa) beyond which all correlations are considered zero.

maxcor (= .99): Maximum allowable value for off-diagonal elements of the obs-obs error correlation matrix. This value is also used to thin the vertical profiles of radiosondes. Used to prevent ill-conditioning of the matrix.

mu (determined from latitude): Correlation between geopotential and streamfunction

nu2 (determined from latitude): Fraction of background error wind variance that is divergent

6.2.2 qc.ctl: parameters controlling the quality control of observations

In the following, the names of the parameters are given, along with their default value and a brief explanation:

bg.crit.dif (= 3): Background check critical difference

do.bgqc (= T): Flag controlling whether to apply the background check

do.dt (= T): Flag controlling whether to apply the data thinning to selected data (usually only applied to satellite data and surface observations).

do.mf (= F): Flag controlling whether to apply the median filter to selected data (usually only applied to satellite data and surface observations).

do.oi (= F): Flag controlling whether to apply the OI QC. The OI QC is presently only implemented in Splus; this flag is ignored unless `tap.execute` is invoked with the `-s` option, which will result in significantly longer execution times.

6.2.3 nes.ctl: parameters controlling the data selection and matrix solution

In the following, the names of the parameters are given, along with their default value and a brief explanation:

d0maxA (computed from maxnhdA and analysis grid spacing): Radius of the analysis volume (expressed as the cosine of the angular great circle distance)

d0maxD (computed from maxnhdA and analysis grid spacing): Radius of the data volume (expressed as the cosine of the angular great circle distance)

d0thin (computed from analysis grid spacing): Thinning radius (expressed as the cosine of the angular great circle distance). Data within this radius of a selected observation are considered for deselection if the maximum number of observations in the data volume is exceeded.

debug (= T): Flag for turning on "debugging" mode. Causes retention of certain work files and generation of diagnostic printouts.

epsrid (= .1): Value added to diagonal of obs-obs correlation matrix in case of ill-conditioning (so-called "ridge addition")

max.iter (= 3): Maximum number of ridge additions

maxnhdA (= 300): Maximum permissible number of headers (horizontal grid points) in analysis volume

maxnhdD (= 75): Maximum permissible number of headers (vertical profiles) in data volume

maxnobA (= F): Maximum permissible number of bodies (analysis values) in analysis volume. Specifying "F" instead of a numeric value disables this criterion for restricting the size of the analysis volume.

maxnobD (= 100): Maximum permissible number of bodies (observation values) in data volume (after subdivision by variable and level; for multivariate analysis, the actual number of bodies may be 2 (u,v) or 3 (z,u,v) times larger). Specifying "F" instead of a numeric value disables this criterion for restricting the size of the data volume.

6.2.4 Other Splus data objects controlling data selection

In the following, the names of the objects are given, along with their default value and a brief explanation:

sellevs (= numeric(0)): Pressure levels (in Pa) for which to read and store radiosonde data. Should be set equal to the mandatory levels if no significant level data is to be processed. Specifying "numeric(0)" means all levels will be used.

data.hdr List with components controlling the ingest and selection of observations:

- **latlim (determined from analysis grid):** Range of latitudes (° N) for which to ingest observations. Not used in AFWA observation ingest routine.
- **lonlim (determined from analysis grid):** Range of longitudes (° E) for which to ingest observations. Not used in AFWA observation ingest routine.
- **maxdelp (=5000):** Maximum vertical separation (in Pa) between observations in data volume and analysis levels.
- **pod.lims (determined from analysis levels and maxdelp):** Range of pressure (in Pa) for which to ingest observations.
- **vars (determined from analysis variables):** Types of variables (expressed in GRIB code numbers) to ingest from observations.

6.3 Customizing analysis grids/variables/levels

As discussed above, this section does not apply to the AFWA implementation of TAP.

The analysis grid, variables, and levels are all specified in a very similar manner: a predefined set of parameter choices, each associated with a unique name, are read from a file at run time. The user then only specifies the name of the desired parameter combination in the user interface. The default set of available variable and level choices are stored in files `d.avars.table` and `d.alevs.table`, respectively. A corresponding file also exists for the analysis grid regions (`d.agrids.table`). The names of these files can not be changed in the user interface.

As was the case for the control files, however, versions of these files in the TAP user's directory are used if they exist, otherwise they are read in from the `$TAPHOME/splus` directory. Customizations can thus be performed for the entire TAP installation (by editing these files in the `$TAPHOME/splus` directory), or for just one TAP user (by editing these files in the TAP user's directory).

6.4 Customizing error statistics

All error statistics are stored in ASCII files. These files are stored in the `$TAPSTATS` directory (see Figure 2). Separate file formats and filename extensions are used for standard deviations of the background and observation error, and for the vertical and horizontal correlation of background and observation errors. As part of the installation of TAP, these files are read in and their content stored in Splus objects in the `$TAPfixed` directory. During execution of the TAP analysis, these Splus objects are used for the calculation of all error statistics.

The default set of error statistics supplied with the initial TAP installation can be modified as follows: a copy of the default set of files should be placed in a separate directory, and modified and/or deleted as desired (deleting a file of horizontal or vertical error correlations is equivalent to assuming uncorrelated errors). The installation of error statistics is then repeated, using the name of the new directory for the environment variable `$TAPSTATS`, and the name of the destination directory for the new Splus objects for the environment variable `$TAPfixed`. Runs of TAP can then use either the default or modified set of statistics by appropriate specification of `$TAPfixed` before the `tap.execute` command.

6.5 Customizing algorithms

It is also possible to substitute user-specific versions of code for those of the default TAP installation.

To substitute individual Splus routines, a modified version of the compiled routine must be placed in the TAP user's `.Data` directory, or the `$TAPHOME/splus/.Data` (if the modified routine is to be used for the entire TAP installation).

Substituting user-modified Fortran or C code for the default TAP versions requires repeating some of the installation steps described in Section 4 and file `install.procedure`, after the modification of the source code. This can again be done for the entire TAP installation (by performing these steps in the TAPHOME directory), or just for one TAP user (by performing these steps in the user's directory). In the latter case, the `-l` switch to the `tap.execute` command must be used.